The Journal of Educational Psychology

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching.

BOARD OF EDITORS:

HAROLD ORDWAY RUGG, *Chairman.*
Teachers College, Columbia University.
Lincoln School of Teachers College.

JAMES CARLETON BELL,
Brooklyn Training School for Teachers.

FRANK NUGENT FREEMAN,
University of Chicago.

ARTHUR IRVING GATES,
Teachers College, Columbia University.

VIVIAN ALLEN CHARLES HENMON,
University of Wisconsin.

RUDOLF PINTNER,
Ohio State University.

LEWIS MADISON TERMAN,
Leland Stanford University.

EDWARD LEE THORNDIKE,
Teachers College, Columbia University.

LAURA ZIRBES, *Assistant Editor.*
Lincoln School of Teachers College.

VOL. XII. No. 1. **JANUARY, 1921** $4.00 a year. 60c. per copy.

CONTENTS

ANNOUNCEMENT OF REORGANIZATION OF THE JOURNAL OF EDUCATIONAL PSYCHOLOGY 1

THE INTERPRETATION AND APPLICATION OF THE INTELLIGENCE QUOTIENT. *Frank N. Freeman* 3

THE CONSTITUTION OF ARITHMETICAL ABILITIES. *Edward L. Thorndike* 14

A CRITICAL STUDY OF THE CONCEPT OF SILENT READING ABILITY. *L. W. Pressey and S. L. Pressey* 25

A COMBINED MENTAL-EDUCATIONAL SURVEY. *Rudolf Pintner and Helen Marshall* 32

DEPARTMENT FOR DISCUSSION OF RESEARCH PROBLEMS. *Laura Zirbes, Editor* 44

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES 50

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION. *A. I. Gates and H. O. Rugg* 52

Published Monthly Except June to August by
WARWICK and YORK, Inc., Baltimore, Md.

Application made for entry as second-class matter at the post office at Baltimore, Md., under the Act of March 3, 1879.

The Journal of Educational Psychology

is published monthly, except June, July and August, by Warwick & York, Inc., 10 E. Center street, Baltimore, Md. The issues of 1921 make Volume XII.

Manuscripts for publication, books or other materials for review and news items should be addressed to Harold Ordway Rugg, Chairman of the Editorial Board, 59 Edgecliff Terrace, Yonkers, N. Y.

The price of the Journal is $4 a year in advance. Foreign postage is 30 cents extra. Single issues are 60 cents each, and less than a full year's subscription is charged at single-issue rate.

Subscribers should notify publishers of changes in address at least four weeks before publication of the issue with which change takes effect.

Educational Psychology Monographs

Studies in the field of educational psychology which are of too great length to be printed in the Journal are issued in book form in the Educational Psychology Monograph Series. These monographs are of varying size and price, and have no definite date for publication. The works listed below have been published, and may be obtained from Warwick & York, Inc., Baltimore, Md.

1. **Moto-Sensory Development.** George V. N. Dearborn. vi, 215 pages, frontispiece. $1.80.
2. **Mental Fatigue.** Max Offner; translated by G. M. Whipple. viii, 133 pages. $1.50.
3. **Backward and Feeble-Minded Children.** Edmund B. Huey. 221 pages, illus. (Out of print.)
4. **When Should a Child Begin School.** W. H. Winch. 108 pages. $1.40.
5. **Spelling Efficiency in Relation to Age, Grade and Sex, and the Question of Transfer.** J. E. Wallace Wallin. vi, 91 pages. $1.40.
6. **Relative Efficiency of Phonetic Alphabets.** Guy Montrose Whipple. 60 pages. 60 cents.
7. **Experimental Studies of Mental Defectives.** J. E. Wallace Wallin. 156 pages, illus. $1.60.
8. **Variations in the Grades of High School Pupils.** Clarence Truman Gray. 120 pages. $1.50.
9. **How I Kept My Baby Well.** Anna G. Noyes. 193 pages and 60 illustrations. $1.80.
10. **The Marking System in Theory and Practice.** I. E. Finkelstein. 88 pages. $1.40.
11. **Inductive Versus Deductive Methods in Teaching.** W. H. Winch. 160 pages. $1.80.
12. **Children's Perceptions.** W. H. Winch. 5-color frontispiece 240 pages. $2.00.
13. **The Psychological Methods of Testing Intelligence.** W. Stern; translated by G. M. Whipple. 160 pages. $1.60.
14. **The Mentality of the Criminal Woman.** Jean Weidensall. 350 pages, illus. $2.50.
15. **The General Value of Visual Sense Training in Children.** Chang Ping Wang. 100 pages. $1.40.
16. **The Doctrine of Formal Discipline in the Light of Experimental Investigation.** Nellie P. Hewins. 120 pages. $1.50.
17. **The Experimental Determination of Mental Discipline in School Studies.** Harold O. Rugg. 150 pages. $1.75.
18. **The Stanford Revision and Extension of the Binet-Simon Intelligence Scale.** Lewis M. Terman and others. 184 pages. $1.80.
19. **The Psychology and Pedagogy of Anger.** Roy Franklin Richardson. 120 pages. $1.40.
20. **The Picture Completion Test.** Rudolf Pintner and Margaret M. Anderson. 112 pages, illus. $1.50.
21. **Deficiency and Delinquency: An Interpretation of Mental Testing.** James Burt Miner. 380 pages. $2.50.
22. **The Influence of Age and Experience on Correlations Concerned With Mental Testing.** Edward Safford Jones. 98 pages. $1.40.
23. **Periodicity With Reference to Mental and Physical Ability.** A. G. Peaks. 120 pages. $1.40. (In process.)
24. **Graphology and the Psychology of Handwriting.** June E. Downey. 180 pages. $1.80.

Warwick & York, Inc., Baltimore, Md.

LIST OF CONTENTS VOLUME XII

J. CARLETON BELL. *Group Tests of Intelligence: An Annotated List* 103

B. R. BUCKINGHAM. *Intelligence and its Measurement: A Symposium*. 271

G. T. BUSWELL. *The Relationship Between Eye Perception and Voice Response in Reading*. 217

FOWLER D. BROOKS. *Rate of Mental Growth, Ages Nine to Fifteen* 501

CLARA F. and LAURA M. CHASSELL. *A Survey of the Three First Grades of the Horace Mann School by Means of Psychological Tests and Teachers' Estimates, and a Statistical Evaluation of the Measures Employed* 72, 243

CECILE COLLOTON AND HAROLD RUGG. *Constancy of the Stanford-Binet I.Q. as shown by Retests* 315

S. S. COLVIN. *Intelligence and its Measurement: A Symposium*. . 136

W. F. DEARBORN. *Intelligence and its Measurement: A Symposium*. 211

F. N. FREEMAN. *The Interpretation and Application of the Intelligent Quotient*. 3

Intelligence and its Measurement: A Symposium. 133

Comments on Professor Peterson's Criticism. 155

The Scientific Evidence on the Handwriting Movement. 253

RAYMOND FRANZEN AND F. B. KNIGHT. *Criteria to Employ in Choice of Tests*. 408

ARTHUR I. GATES. *Educational Psychology at the Chicago Meeting of Scientific Societies* 63

The True-False Test as a Measure of Achievement in College Courses. 276

An Experimental and Statistical Study of Reading and Reading Tests 303, 378, 445

ARTHUR I. GATES, CHAS. H. JUDD AND LAURA ZIRBES. *Special Review of Mrs. Burgess' Monograph on Silent Reading* . . . 347

M. E. HAGGERTY. *Intelligence and Its Measurement: A Symposium*. 212

V. A. C. HENMON. *An Experimental Study of the Value of Word Study*. 98

Intelligence and its Measurement: A Symposium 195

J. P. HERRING. *Verbal and Abstract Elements in Intelligence Examinations* 511

CHAS. H. JUDD, ARTHUR I. GATES AND LAURA ZIRBES. *Special Review of Mrs. Burgess' Monograph on Silent Reading* . . . 347

T. L. KELLEY. *Transmutation of Values on the Thorndike and Ayres Handwriting Scales: A Correction*. 288

F. B. KNIGHT AND RAYMOND FRANZEN. *Criteria to Employ in Choice of Tests*. 408

E. E. LINDSAY. *Personal Judgments* 413

418633

HELEN MARSHALL AND RUDOLF PINTNER. *A Combined Mental-Educational Survey* 23
Results of the Combined Mental-Educational Survey Tests . . . 82
LEIGH MUDGE. *Time and Accuracy as Related to Mental Tests* . 159
GARRY C. MYERS. *Prophecy of Learning Progress by Beta* . . . 228
L. A. PECKSTEIN. *Masses vs. Distributed Effort in Learning* . . . 92
JOSEPH PETERSON. *The Growth of Intelligence and the Intelligence Quotient* 148
Intelligence and its Measurement: A Symposium 98
RUDOLF PINTNER AND HELEN MARSHALL. *A Combined Mental-Educational Survey* 32
Results of the Combined Mental-Educational Survey Tests . . . 82
RUDOLF PINTNER. *Intelligence and its Measurement: A Symposium* 139
LOUISE E. POULL. *Constancy of I.Q. in Mental Defective, According to the Stanford Revision of Binet Tests* 323
L. W. AND S. L. PRESSY. *A Critical Study of the Concept of Silent Reading Ability* 25
S. L. PRESSY. *Intelligence and its Measurement: A Symposium* 144
HAROLD RUGG AND CECILE COLLOTON. *Constancy of the Stanford-Binet I.Q. as shown by Retests* 315
HAROLD RUGG. *Is the Rating of Human Character Practicable?* 425, 485
B. RUML. *Intelligence and Its Measurement: A Symposium* . . . 143
LEWIS M. TERMAN. *Mental Growth and the I.Q* 325, 401
Intelligence and Its Measurement: A Symposium 127
PAUL W. TERRY. *The Reading Problem in Arithmetic* 365
E. L. THORNDIKE. *The Constitution of Arithmetical Abilities* . . 14
Intelligence and its Measurement: A Symposium 124
The Psychology of Drill in Arithmetic: The Amount of Practice 183
L. L. THURSTONE. *Intelligence and its Measurement: A Symposium* 202
HERBERT WOODROW. *Intelligence and its Measurement: A Symposium* 207
J. E. WALLACE WALLINS. *The Results of Retests by Means of the Binet Scale* 392
ANGELINA WEEKS. *The Terman Vocabulary as a Group Test* . . 531
J. WYMAN AND MIRIAM WENDLE. *What is Reading Ability?* . . 517
LAURA ZIRBES, ARTHUR I. GATES AND CHAS. H. JUDD. *Special Review of Mrs. Burgess' Monograph on Silent Reading* . . . 347

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY

Vol. XII JANUARY, 1921 No. 1

ANNOUNCEMENT OF THE REORGANIZATION OF THE JOURNAL OF EDUCATIONAL PSYCHOLOGY.

With this issue the Editorial Board of the JOURNAL OF EDUCATIONAL PSYCHOLOGY is reorganized and enlarged. The new board is now engaged in formulating the policy of the Journal and in outlining its contents for the year 1921. The details of this policy and of the outline of contents will be announced in the next issue. We wish now to comment briefly upon the broader outline of that policy.

The Journal will be devoted primarily to the scientific study of problems of learning and teaching. The editors' chief purpose is to make it a clearing house for the discussion of scientific investigation and experimentation, which is necessary to improve permanently teaching in the public schools. They will attempt to fulfill this purpose by organizing series of articles, by creating discussion departments, and by making the Journal a prompt and complete clearing house for the review of pertinent educational publications in the field. To do that they are now securing series of articles by leading workers in educational psychology and educational experimentation on such matters as:

1. Desirable changes in public school curricula from the standpoint of experimentation and investigation of learning. Professor Thorndike's article on "The Constitution of Arithmetical Abilities" in the present issue is the first article of this series. It will be followed by others written by leaders who have investigated the learning and teaching of writing, spelling, reading, history, geography, civics, home economics and high school mathematics. Thus the Journal should appeal to public school supervisors and pro-

gressive teachers, as well as to specialists in educational psychology.

2. Critical and evaluative articles on the construction and use of intelligence and educational tests. The JOURNAL OF EDUCATIONAL PSYCHOLOGY was one of the first publication agencies to commit itself definitely to the support and improvement of the intelligence test movement. The editors have an important interest in the continuation of that leadership by the Journal.

3. Articles and discussions of the content of courses in educational psychology which will appeal to normal school teachers of the subject. The editors wish to contribute as much as possible to the improvement of the training of public school teachers and believe that the application to normal school curricula of current scientific thinking in educational psychology will go far towards bringing about this result.

4. A new department for discussion of research problems has been organized, and will be conducted by Miss Laura Zirbes. The first publication of material from this department appears in the present issue.

5. Each month the Journal will attempt to acquaint its readers with the articles in educational psychology that have appeared in the current issue of other magazines by means of brief annotated references.

6. The editors wish to stimulate the publication of important miscellaneous findings from investigation and experimentation. Beginning with the February issue, they will present, in short discussions, statistical findings and concise interpretations from investigations of learning, intelligence and educational testing, correlation work and the like.

7. The editors will make a serious attempt to review in each issue of the Journal books, monographs and other types of important contribution in educational psychology and closely related fields of education which have appeared within the previous month or two.

THE INTERPRETATION AND APPLICATION OF THE INTELLIGENCE QUOTIENT.

FRANK N. FREEMAN,
University of Chicago.

The purpose of this paper is to discuss the relationship between the IQ as a measure of the mental capacity of the individual and the facts of mental development. The IQ was first used as a means of expressing the relative mental ability of individuals as measured by the Binet scale. The first measure which was used with this scale was *mental age.* This, however, is an absolute and not a relative measure. The first relative measure used was the difference between the mental age and the chronological age, which may be called the *mental age difference.* This expresses the individual's superiority or inferiority in terms of a year's mental growth as a unit. It was soon discovered, however, that this measure, when applied to the Binet scale, would result in an apparent greater retardation or acceleration in the case of older children than in the case of younger ones.

The fundamental explanation of this fact was found in the greater overlapping of the scores of older children than of younger ones. To give a specific illustration, in the case of a number of tests which were used in the Binet scale and which were found standard for the fifth year, about twenty-three per cent more six year old than five year old children passed the tests. When the comparison was made with the tests for ten year olds, however, it was found that not until the twelfth year was reached did the older group excel the younger group by twenty-three per cent. Another way of putting the matter is to say that the five year old group overlapped the six year old group to the same degree as the ten year old group overlapped the twelve year old group. As a consequence of this fact, a six year old pupil whose score equaled that of the median of the five year old child was found to occupy about the same position in the distribution of his group as was occupied by the twelve year old child whose score equaled that of a ten year old. The child's mental capacity could be represented by a ratio between his mental age and his chronological age and this ratio would remain constant. To put it in another way, the use of such a ratio gives a distribution of indices of ability which have approximately the

same range from year to year. This make the measures comparable throughout the period of the individual's mental growth.

We have been dealing thus far with facts of observation and with their use empirically to get measures of ability which shall be comparable from year to year and which will fit a particular scale. Thus far, no theory of mental development has been proposed. The point with which we are next concerned is the explanation of the fact that the IQ is the measure of ability which fits the Binet scale, and following this will come the question of the application of these facts and the consequent theories to the methods of expressing the results of other scales. To be specific, the IQ, or the ratio of mental age to chronological age, has been used in connection with certain point scales. The justification for this use will be one of the subjects of our inquiry.

It may be noted, parenthetically, that we are not dealing here with the ratio used by Yerkes in his point scale, namely, the coefficient of intellectual ability. This cofficient is calculated by finding the ratio between the individual's score and the median score for his age. We may, however, in connection with the other discussion, consider the implication of this index also.

What now is the nature of intellectual ability which is implied by the use of the IQ? The most common assumption concerning this matter is that the rate of intellectual growth is not uniform but regularly decreases with advancing age. The curve of mental growth which may be taken to satisfy roughly the requirement of the IQ is given in Woodrow's "Brightness and Dullness of Children." See Figure 1. This curve is apparently drawn somewhat empirically. The form of the curve is not in this case the sole condition of the validity of the IQ, but it is the chief one. In addition to a decreasing rate of mental growth we find here represented also some divergence between the lines of growth which represent different levels of ability.

In order to illustrate the way in which the IQ works out at various ages, and remains constant at successive age for a given level of ability, four cases of individuals on the moron level have been represented at the chronological age 3, 6, 9 and 15. The mental ages are indicated by the points on the normal curve at which it is cut by the horizontal lines drawn from the points on the curve for the moron

level which correspond to the respective chronological ages. The divergence of the two lines is indicated by the increasing length of the vertical lines which join them at the respective ages. The decrease in the IQ at age 15 indicates that the lines diverge too much at this part or that they approach the horizontal too quickly.

It may be shown mathematically that if the rate of growth were the only factor determining the IQ, and if the IQ were valid, the curve would be logarithmic, the formula being $y = \log x$. This is illustrated in Figure 2. Assuming that the growth curves of all individuals, including superior and inferior ones, followed this form and were equidistant from one another on the vertical axis, this form of development would produce the degree of overlapping in successive years implied in the IQ. Compare for example the ratio of mental age to chronological age in the case of illustrative individuals of the chronological ages, 3, 6, 9 and 15, and mental ages, 2, 4, 6 and 10, respectively.

This, however, is only one of two possible conditions which may alone produce the increasing overlapping necessary to make valid the IQ. The second possible factor is illustrated in Figure 3. We have here represented a form of mental growth which follows a straight line. The lines of development of different individuals, however, do not run parallel to one another, but diverge. Thus the range of distribution at successive years increases proportionately to the size of the score. The way in which this supports the validity of the IQ is illustrated again by the relation of mental age to chronological age of inferior children at the ages of 3, 6, 9 and 15. These two factors might, of course, be combined in various proportions.

The Binet scale is so constituted that it gives no indication whether the one or the other of these two assumptions is the correct one or whether there exists in fact a combination of them. As already remarked, Woodrow assumes some divergence on growth curves as well as a decrease in rate of growth. The Binet scale does not give an individual's intelligence rating in terms of a scale which is independent of the age levels themselves. Mental rating is always in terms of abilities at different age levels and therefore it is impossible to determine the nature of growth except in age terms.

In order to get a further line upon this problem, it is necessary to examine the age norms in those tests which use the same scale

Fig. 1. Growth curves representing developmental and divergent deviation as a basis for the intelligence quotient (Terman)

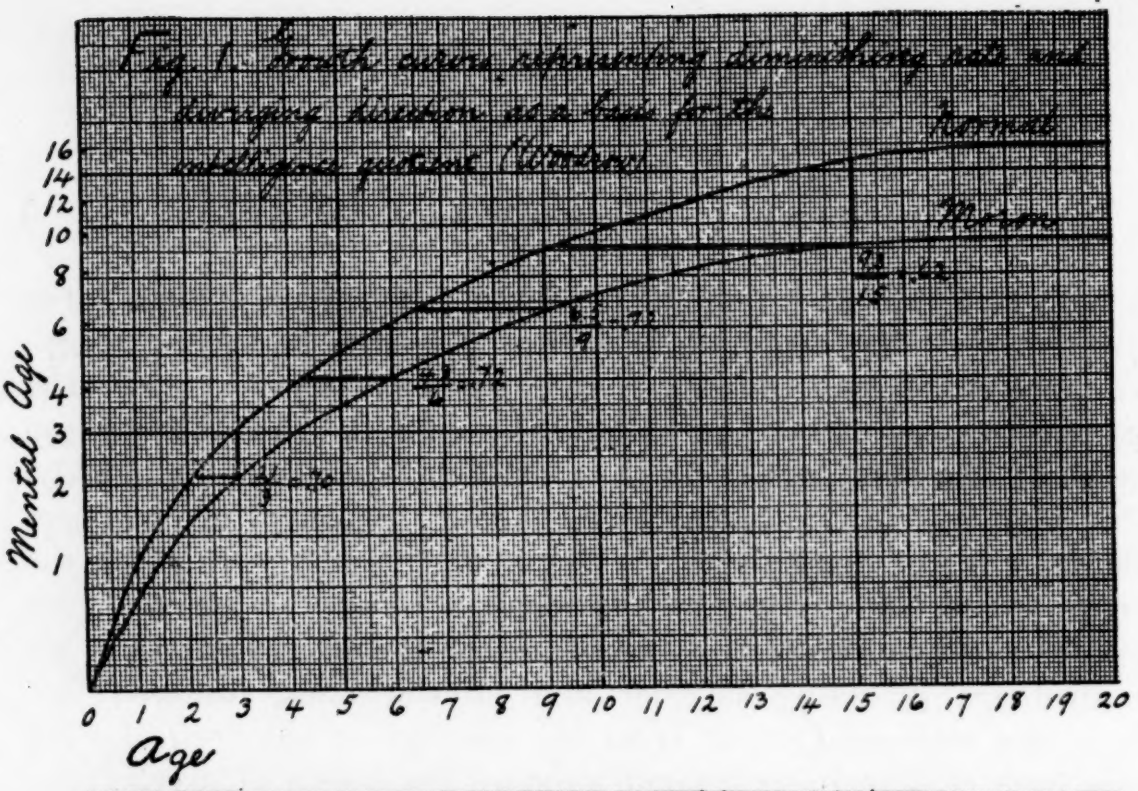


Fig. 2. Illustration of the curve $y = \log x$ as the basis for the intelligence quotient

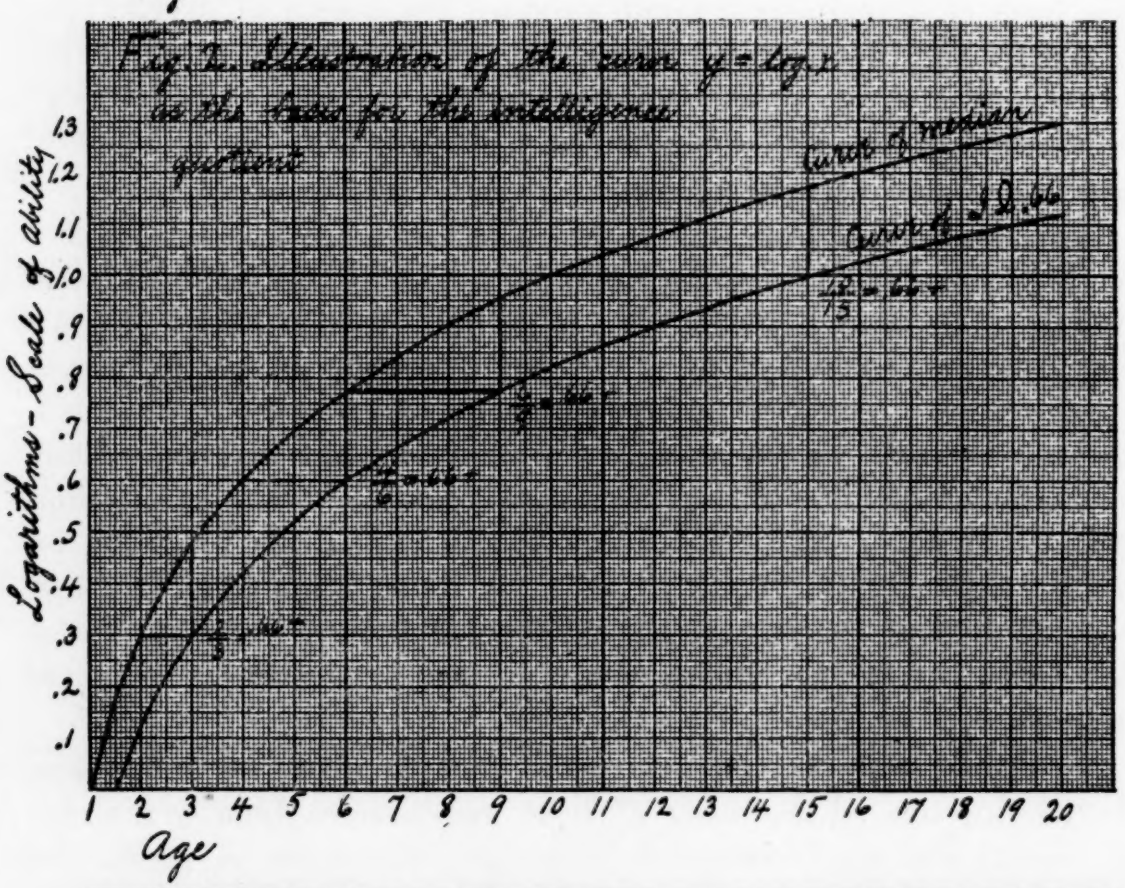
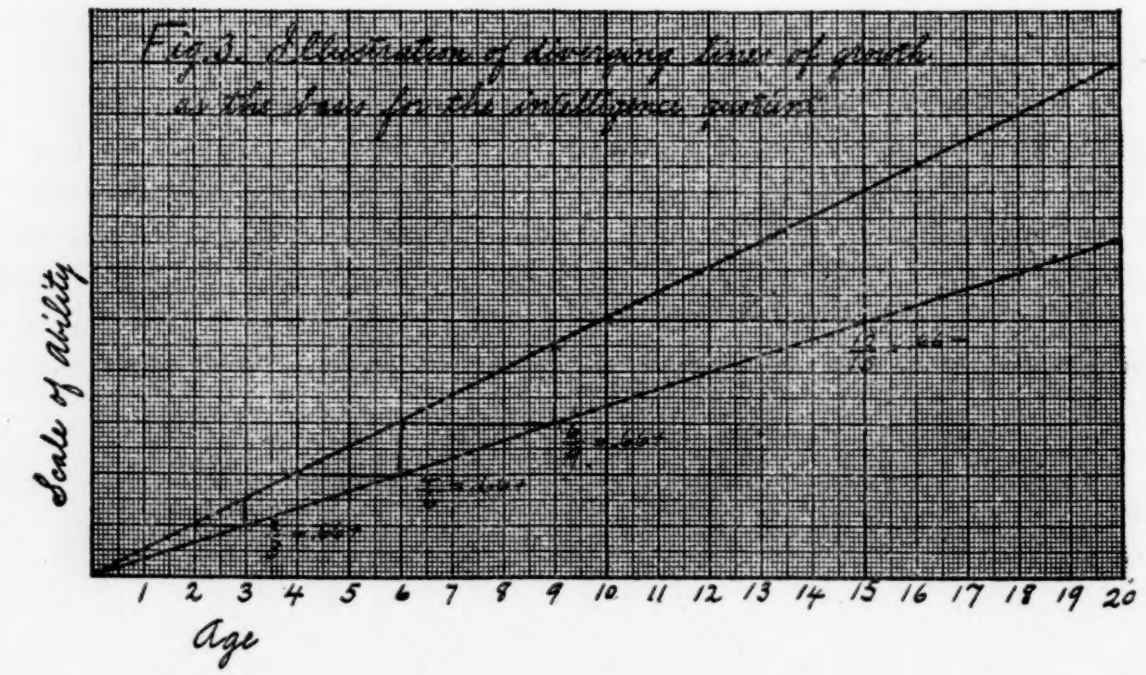


Fig. 3. Illustration of diverging laws of growth as the basis for the intelligence quotient



throughout the ages which are to be compared. This is true of the various point scales which are now in common use.

We may first inquire whether the curve of mental growth follows or approaches the logarithmic curve, or in more general terms what its form is. We have been given age norms for several point scales. The validity of these norms may be questioned beyond the age of thirteen or fourteen, on the ground of a selection of cases beyond these years. We are safe, therefore, only in reference to the years below this.

The first point scale for which we have norms is the one devised by Yerkes. The curve is shown in Figure 4. A casual glance at this figure might give the impression that it represents a decreasing rate of growth. More careful inspection, however, shows that it is made up of two parts, one for the years four to twelve, and the other for the later years. The line of the norms from four to twelve follows almost exactly a straight line. Since this is the most reliable part of the curve, we need not consider the rest.

The second growth curve, based on a large number of children tested by a point scale, is that of Pressey, as shown in Figure 5. This gives very nearly a straight line curve also, extending to age sixteen, although there is a slight divergence covering years eleven to fourteen. If we extend the examination only to years eight to fourteen, the line of development is practically straight. Pressey gives also the norms for this cross-out test scale, beginning with age four. The curve for these is shown in Figure 6. This gives practically a straight line growth up to and including age nine. At ages ten and eleven we are apparently approaching, as in many other cases, the limits of difficulty of the tests.

Figure 7 shows the same facts for the Otis Test, Figure 8 for the Haggerty Tests, Figure 9 for the Pintner Non-language Tests, Figure 10 for the National Intelligence Test, and Figure 11 for the Illinois Examination. In all of these but the Pintner and the National Intelligence Test, the line of growth is almost exactly straight. In fact, the norms which Otis derives from his data, represented by the light lines in Figure 7, follow exactly a straight line up to age fourteen. This may be compared with the actual growth curves which are represented by the heavy lines and which were derived from the table of the distribution scores.

It is apparent that the assumption that intellectual growth follows a curve which approaches the logarithmic curve is not borne

Fig. 8. Norms for Haggerty's Delta 1 and Delta 2

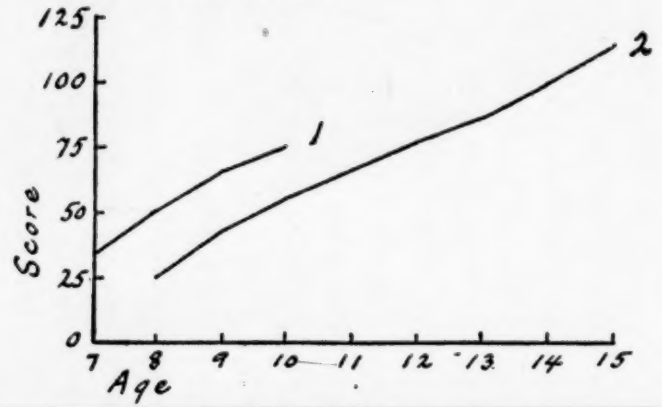


Fig. 9. Curves for three percentile groups in Pintner's non-language test

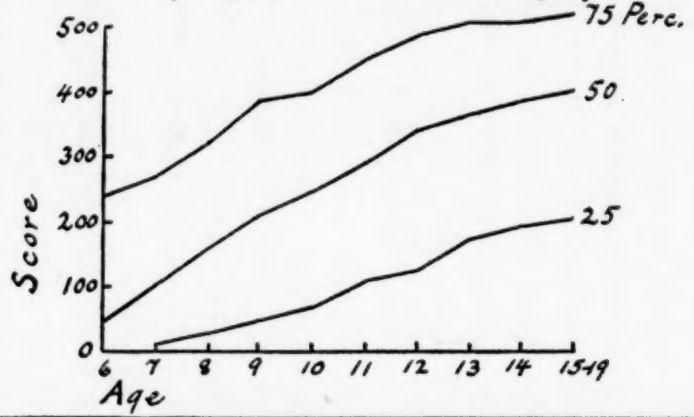


Fig. 10. Provisional age standards, National Intelligence Tests

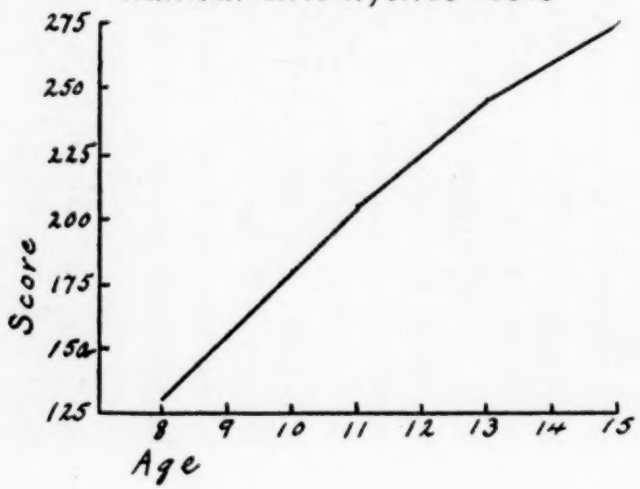
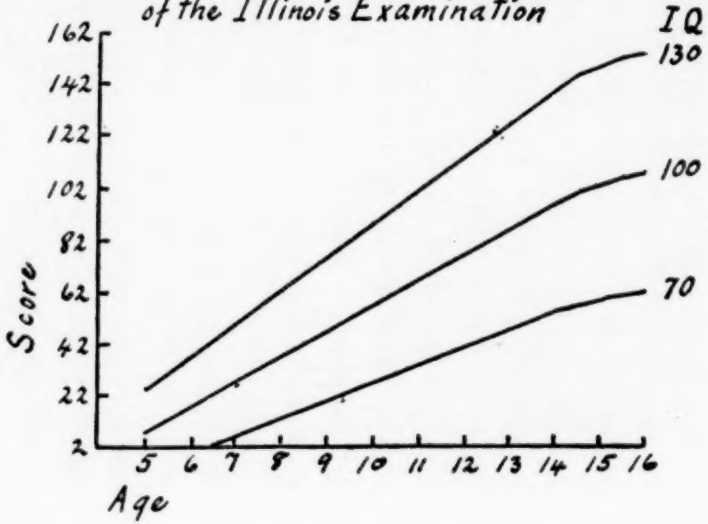


Fig. 11. Three IQ age lines from norms of the Illinois Examination



out by the results of the point scale examinations. There are two conditions which render the scope and the accuracy of the interpretation of the results somewhat doubtful. The first is a hypothetical slowing up of the rate of growth somewhere in the adolescent period. The second, which complicates and renders difficult of interpretation the facts in regard to this retardation in growth, grows out of the nature of the tests themselves. It is quite apparent from a number of the test norms that there may be an apparent retardation in rate of growth which is due solely to the organization of the test itself. See for example, the breaks in the curves shown in Figures 6 and 8 (curve 1) which come before the adolescent period and at ages when the growth is shown by other tests to be proceeding at an undiminished rate. It is, therefore, a hazardous proceeding to assume that the rate of growth diminishes at any particular point on the ground of the results from one test alone. Only a still more extensive investigation than has thus far been made will indicate the point at which the rate does diminish and the degree of diminution which comes at this and later ages. The preponderance of evidence, however, seems to indicate that up to some age in early adolescence at least the rate of growth is approximately uniform.

A test which represented in its norms this type of growth, then, could not legitimately use the IQ, unless the other condition were present, namely, a divergence of lines of growth of individuals at different levels or, in other words, an increase in the range of distribution at succeeding ages. We may then examine this matter in the case of those tests for which the distribution at each age is given. The measures of distribution have not been given for all the tests for which norms have been furnished. We have them, however—the Pressey cross-out test (Figure 6), the Otis test (Figure 7), and the Pintner non-language test (Figure 9). The interquartile range has also been calculated for the Chicago Intelligence Test at the various grades. It is as follows:

Grade.....	VII	VIII	IX	X	XI	XII
Range	15.0	11.5	17.6	14.9	16.0	15.9.

In the case of the Illinois Examination, a table is given for the calculation of an individual's IQ which implies a regular and uniform divergence in the scores at succeeding ages. This table is doubtless based upon data from the results of the examination, but

Fig. 4. Age medians in the Yerkes Scale

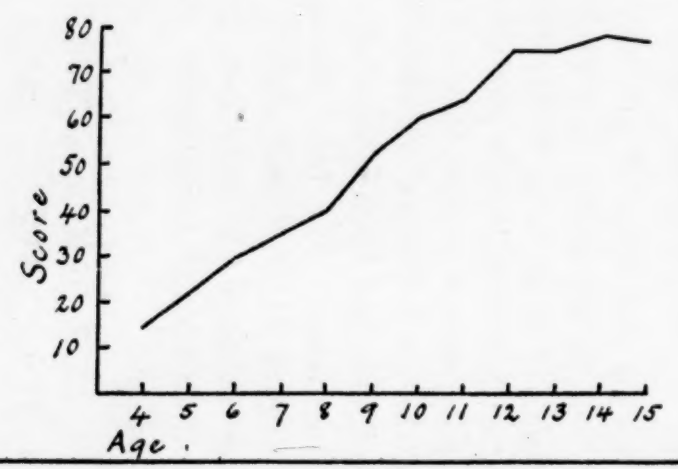


Fig. 5. Age medians for boys in the Pressey Mental Survey Scale No 1

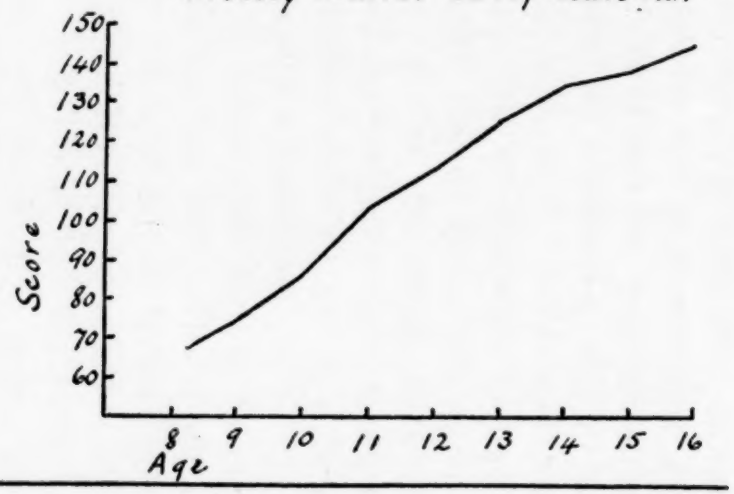


Fig. 6. Curves for the various percentile groups in the Pressey Cross-out Test

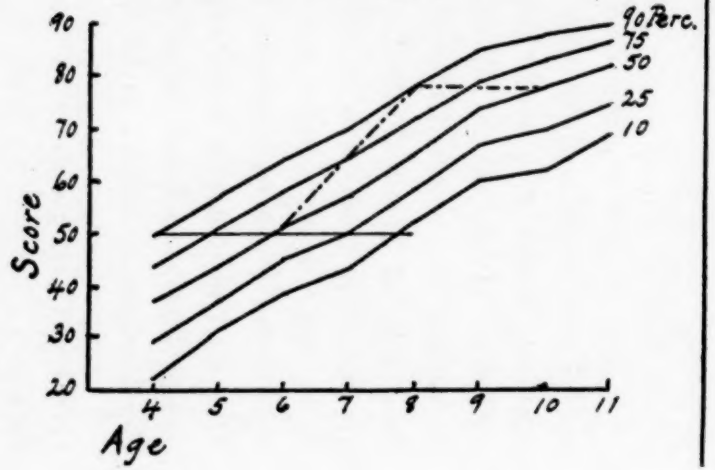
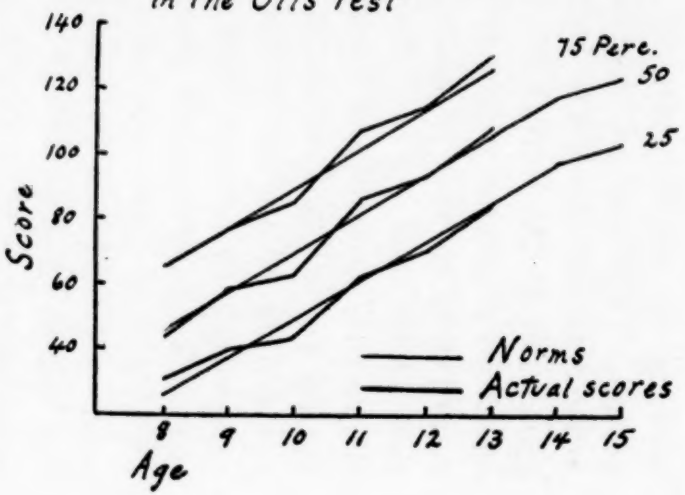


Fig. 7. Curves for three percentile groups in the Otis Test



as these data are not supplied in their original form we shall confine our attention to those cases in which we have the original scores.

Of the four examples which we have before us, which are based on an adequate number of cases, three of them indicate parallel lines of growth, namely, the Pressey test, the Otis test, and the Chicago test. In the Pintner test, the seventy-five percentile curve approaches more closely to the median curve at the higher ages. The twenty-five percentile curve, however, diverges from the median up to age twelve and then remains parallel. It should be noted that the twenty-five percentile curve comes very close to the zero score in the lower ages and this limit in possible scores may be responsible for the fact that it approaches the median at these points. In fact it would be impossible to make a score below year ten which diverges as widely from the median as does the score of the twenty-five percentile at the year twelve, and at the year ten the score would be only slightly over twenty points out of a possible 500 plus. It appears likely therefore that the character of the distribution and possibly the course of the median scores in the early ages is due to the difficulty of the test for the younger children.

The composite effect of the various apparently rather fortuitous factors in the determination of the curves from the Pintner test may be indicated by the following measure of overlapping. We may compare the median score for each age with the percentile scores of the ages next above and below it. If these percentiles approach the fifty percentile (median) at the higher ages the overlapping increases. The facts are shown as follows:

Percentile of age higher or lower which corresponds to the median score at successive ages.

Age	6	7	8	9	10	11	12	13	14
Age higher.	40	42	45	45	43	43	46	47	47
Age lower.	63	59	57	55	57	57	54	53	53

It may be seen that there is very slight evidence of increasing overlapping from ages 7 to 11. The score for age 6 is too erratic to be relied on and the scores above age 12 follow a different course than do those below.

It may be remarked, parenthetically, that Yerkes' *cofficient of intelligence* implies an increase in the range of distribution with

advancing age, and if this increase does not occur this index is invalid.

It appears from these facts that both of the assumptions which may serve to explain the validity of the IQ in the case of the Binet scale are in question. There seems to be evidence of considerable weight that the typical intellectual growth follows a uniform rate, within at least the period covered roughly by the elementary school. We are here, of course, assuming that the measure of mental ability is a composite one, which gives opportunity for the exercise of a variety of functions, and for a wide range of scores ranging from considerably below the ability of the youngest children tested to considerably above the ability of the oldest. The evidence seems further to indicate that the individuals above and below the median develop at about the same rate, though in different levels as compared with the median child. At any rate, the correctness of the usual assumptions is called in question by the facts before us.

This last tentative conclusion may receive further comment in view of practical experience which seems to contradict it. Our everyday observation seems to indicate that the superior child advances more rapidly than the median child and the inferior child more slowly. In many experiments, for example, it has been shown that the upper fifth, or thereabouts, of a grade may be segregated, and may make much more rapid progress than the mass of the children who remain within the ordinary grade. This, however, may be susceptible of an entirely different explanation than that the superior children are developing with unusual rapidity.

An illustration may be taken from Figure 6. Assume that a group of ninety percentile children, which represents the median of the upper fifth, started in school at the age of six with a group of children of median ability. The rate of advancement in school of these median children might be represented by the dot and dash line. It will be seen in the figure that this line meets the ninety percentile line at eight years, meaning that this group of children has, by the time they are eight years old, and after two years' schooling, reached a degree of school attainment which corresponds to their native ability. This school attainment is seen to be equivalent to that which the median children have reached at ten years, after four years' training.

In other words, these superior children have, as in the experiments alluded to, made four years' advancement in two. This is

not, however, because their mental growth has been twice as rapid as the others, but because in the space of two years' time they have risen to their mental level in school work. A similar illustration might be drawn from the inferior children, who, by retardation and repeating grades, sink to their own mental level.

This is the condition which results from the entrance of children of all mental levels at the same age and from giving them all the same grade of work. If, on the contrary, all children entered school at the same mental level, which is represented by the horizontal line opposite score 50, they would enter at widely different ages, the ninety percentile children entering approximately at four, the seventy-five percetile at five, the median at six, the twenty-five percentile at seven, and the ten percentile at eight. Assuming that this were done and that they were all given the same curriculum, the data before us would indicate that they would advance at equal rates, up to the period at least of the retardation or cessation of mental growth. Assuming that this cessation came at the same age for the different groups, it would come at different levels in school attainment, the ninety percentile child being able to continue his advancement four grades farther than the ten percentile child.

The assumption is not made that this conclusion is strictly accurate. It indicates, however, that it is quite possible to interpret the main facts of our ordinary observation on the basis of the assumption of straight line development, and parallel development at different levels, and this strengthens the conclusion that the assumption that decreasing rate and diverging lines represent the nature of mental growth must be called into serious question.

Our discussion started with the concrete problem of the IQ and with the assumptions which underly it. It has led us to conclude that the application of the IQ to other than the Binet scale must be made with great caution and only after determining that it is a suitable method of representing the scores in other tests. We have been led beyond this practical issue to the consideration of the general nature of mental development and have been led to call in question the time honored conceptions in regard to it. The evidence at hand is sufficient to raise serious doubts about these conceptions and to stimulate to further inquiry to establish on firmer foundation the general theories concerning mental growth.

THE CONSTITUTION OF ARITHMETICAL ABILITIES

EDWARD L. THORNDIKE,
Teachers College, Columbia University.

When the analysis of the mental functions involved in arithmetical learning is made thorough it turns into the question, 'What are the elementary bonds or connections that constitute these functions?'; and when the problem of teaching arithmetic is regarded, as it should be in the light of present psychology, as a problem in the development of a heirarchy of intellectual habits, it becomes in large measure a problem of the choice of the bonds to be formed and of the discovery of the best order in which to form them and the best means of forming each in that order.

The Importance of Habit Formation.—The importance of habit formation or connection-making has been grossly underestimated by the majority of teachers and writers of text-books. For, in the first place, mastery by deductive reasoning of such matters as 'carrying' in addition, 'borrowing' in subtraction, the value of the digits in the partial products in multiplication, the manipulation of the figures in division, the placing of the decimal point after multiplication or division with decimals, or the manipulation of the figures in the multiplication and division of fractions is impossible or extremely unlikely in the case of children of the ages and experience in question. They do not as a rule deduce the method of manipulation from their knowledge of decimal notation. Rather they learn about decimal notation by carrying, borrowing, writing the last figure of each partial product under the multiplier which gives that product, etc., etc. They learn the method of manipulating numbers by seeing them employed, and by more or less blindly acquiring them as associative habits.

In the second place, we, who have already formed and long used the right habits and are thereby protected against the casual misleadings of unfortunate mental connections, can hardly realize the force of mere association. When a child writes sixteen as 61, or finds 428 as the sum of 15, or gives 642 as an answer to

19
16
18
—

27×36, or says that 4 divided by $\frac{1}{4} = 1$, we are tempted to consider him mentally perverse, forgetting or perhaps never having understood that he goes wrong for exactly the same general reason that we go right, namely, the general law of habit formation. If we study the cases of 61 for 16, we shall find them occurring in the work of pupils who, after having been drilled in writing 26, 36, 46, 62, 63, and so on, in which the order of the six in writing is the same as it is in speech, return to writing the 'teen number. If our language said onety-one for eleven and onety-six for sixteen, we should probably never find such errors except as 'lapses' or as the results of misperception or lack of memory. They would then be more frequent *before* the 20's, 30's, etc., were learned.

If pupils are given much drill on written single column addition involving the higher decades (each time writing the two-figure sum), they are forming a habit of writing 28 after the sum of 8 6 9 and 5 is reached, and it should not surprise us if the pupil still occasionally writes the two figure sum for the first column though a second column is to be added also. On the contrary, unless some counter force influences him, he is absolutely sure to make this mistake.

In connection with the third mistake quoted there is opportunity for a very instructive experiment, namely, to give to a group of children who are just to be taught 'long' multiplication a set of examples where say a three place number is to be multiplied by a three place number, asking them to multiply. Some will do nothing. Some will subtract. Some will multiply them as shown above. Some will multiply the upper number by one or more figures of the lower number. All will exemplify the action of the law of association. They do what they have done in the situation most like the present. Since they have rarely added, never multiplied or divided and have often subtracted with two three place numbers, some will be led by habit to neglect the request to multiply. Since they have multiplied 7 by 6 getting 42, and 3 and 2 getting 6, and 4 by 5 getting 20, some will write

$$\begin{array}{r} 26 \\ \hline 642 \end{array}$$

and so on. It is with or against the force of such habits as these that the new habits of 'long' multiplication are formed.

The last mistake quoted ($4 \div \frac{1}{4} = 1$) is interesting because here we have possibly one of the cases where deduction from psychology alone can give constructive aid to teaching. Multiplication and division by fractions have been notorious for their difficulty. The former is now alleviated by using *of* instead of X until the new habit is fixed. The latter is still approached with elaborate caution and with various means of showing why one must 'invert and multiply' or 'multiply by the reciprocal'.

But in the author's opinion it seems clear that the difficulty in multiplying and dividing by a fraction was not that children felt any logical objections to concelling or inverting. I fancy that the majority of them would cheerfully invert any fraction three times over or cancel numbers at random in a column if they were shown how to do so. But if you are a youngster inexperienced in numerical abstractions and if you have had *divide* connected with 'make smaller' three thousand times and never once connected with 'make bigger', you are sure to be somewhat impelled to make the number smaller the three thousand and first time you are asked to divide it. Some of my readers will probably confess that even now they feel a slight irritation or doubt in saying or writing that $16 \div \frac{1}{8} = 128$.

The habits that have been confirmed by every multiplication and division by integers are, in this particular of *"the ratio of result to number operated upon"*, directly opposed to the formation of the habits required with fractions. And that is, I believe, the main cause of the difficulty. Its treatment then becomes easy, as will be shown later.

These illustrations could be added to almost indefinitely, especially in the case of the responses made to the so-called 'catch' problems. The fact is that the learner rarely can, and almost never does, survey and analyze an arithmetical situation and justify what he is going to do by articulate deductions from principles. He usually feels the situation more or less vaguely and responds to it as he has responded to it or some situation like it in the past. Arithmetic is to him not a logical doctrine which he applies to various special instances, but a set of rather specialized habits of behavior toward certain sorts of quantities and relations. And in so far as he does come to know the doctrine it is chiefly by doing the will of the master. This is true even with the clearest exposi-

tions, the wisest use of objective aids and full encouragement of originality on the pupil's part.

Lest the last few paragraphs be misunderstood, I hasten to add that the psychologists of today do not wish to make the learning of arithmetic a mere matter of acquiring thousands of disconnected habits, nor to decrease by one jot the pupil's genuine comprehension of its general truths. They wish him to reason not less than he has in the past, but more. They find, however, that you do not secure reasoning in a pupil by demanding it, and that his learning of a general truth without the proper development of organized habits back of it is likely to be, not a rational learning of that general truth, but only a mechanical memorizing of a verbal statement of it. They have come to know that reasoning is not a magic force working in independence of ordinary habits of thought, but an organization and coöperation of those very habits on a higher level.

The older pedagogy of arithmetic stated a general law or truth or principle, ordered the pupil to learn it, and gave him tasks to do which he could not do profitably unless he understood the principle. It left him to build up himself the particular habits needed to give him understanding and mastery of the principle. The newer pedagogy is careful to help him build up these connections or bonds ahead of and along with the general truth or principle, so that he can understand it better. The older pedagogy commanded the pupil to reason and let him suffer the penalty of small profit from the work if he did not. The newer provides instructive experiences with numbers which will stimulate the pupil to reason so far as he has the capacity, but will still be profitable to him in concrete knowledge and skill, even if he lacks the ability to develop the experiences into a general understanding of the principles of numbers. The newer pedagogy secures more reasoning in reality by not pretending to secure so much.

The newer pedagogy of arithmetic, then, scrutinizes every element of knowledge, every connection made in the mind of the learner, so as to choose those which provide the most instructive experiences, those which will grow together into an orderly, rational system of thinking about numbers and quantitative facts. It is not enough for a problem to be a test of understanding of a principle; it must also be helpful in and of itself. It is not enough for an

example to be a case of some rule; it must help review and consolidate habits already acquired or lead up to and facilitate habits to be acquired. Every detail of the pupil's work must do the maximum service in arithmetical learning.

It is then profitable to consider detailed cases of bonds now often neglected which deserve attention. Those listed below are among the most important.

Numbers As Measures of Continuous Quantities.—The numbers one, two, three, 1, 2, 3, etc., should be connected soon after the beginning of arithmetic each with the appropriate amount of some continuous quantity like length or volume of weight, as well as with the appropriate sized collection of apples, counters, blocks and the like. Lines should be labelled 1 foot, 2 feet, 3 feet, etc.; one inch, two inches, three inches, etc.; weight should be lifted and called one pound, two pounds, etc.; things should be measured in glassfuls, handfuls, pints, and quarts. Otherwise the pupil is likely to limit the meaning of, say, four to four sensibly discrete things and to have difficulty in multiplication and division. Measuring or counting by insensibly marked off repetitions of a unit, binds each number name to its meaning as—times *whatever* 1 is, more surely than mere counting of the units in a collection can, and should reinforce the latter.

(2) *Additions in the Higher Decades.*—In the case of all save the very gifted children, the additions with higher decades—that is, the bonds, $16 + 7 = 23$, $26 + 7 = 33$, $36 + 7 = 43$, $14 + 8 = 22$, $24 + 8 = 32$, and the like—need to be specifically practiced until the tendency becomes generalized. 'Counting' by 2s beginning with 1, and with 2, counting by 3s beginning with 1, with 2, and with 3, counting by 4s beginning with 1, with 2, with 3, and with 4, and so on, make easy beginnings in the formation of the decade connections. Practice with isolated bonds should soon be added to get freer use of the bonds. The work of column addition should be checked for accuracy so that a pupil will continually get beneficial practice rather than 'practice in error'.

(3) *The Uneven Divisions.*—The quotients with the remainder for the divisions of every number to 19 by 2, every number to 29 by 3, every number to 39 by 4, and so on should be taught as well as the even divisions. A table like the following will be found a convenient means of making these connections:

10 = . . . 2s.
10 = . . . 3s and . . . rem.
19 = . . . 4s and . . . rem.
10 = . . . 5s.
11 = . . . 2s and . . . rem.
11 = . . . 3s and . . . rem.
.
.
.
89 = . . . 9s and . . . rem.

These bonds must be formed before short division can be efficient, are useful as a partial help toward selection of the proper quotient figures in long division, and are the chief instruments for one of the important problem series in applied arithmetic,—"How many X's can I buy for y cents at z cents per x and how much will I have left?" That these bonds are at present sadly neglected is shown by Kirby ('13) who found that pupils in the last half of grade 3 and the first half of grade 4 could do only about four such examples per minute (in a ten-minute test), and even at that rate made far from perfect records, though they had been taught the regular division tables. Sixty minutes of practice resulted in a gain of nearly 75 per cent in number done per minute, with an increase in accuracy as well.

(4) *The Equation Form.*—The equation form with an unknown quantity to be determined, or a missing number to be found should be connected with its meaning and with the problem attitude long before a pupil begins algebra, and in the minds of pupils who never will study algebra.

Children who have just barely learned to add and subtract learn easily to do such work as the following:

4 + 8 = . . .
5 + . . . = 4.
. . . + 3 = 11.
. . . = 5 + 2.
16 = 7 + . . .
12 = . . . + 5.

The equation form is the simplest uniform yet devised to state a quantitative issue. It is capable of indefinite extension if certain easily understood conventions about parenthesis and fraction signs are learned. It should be employed widely in accounting and the treatment of commercial problems, and would be except for outworn conventions. It is the chief contribution of algebra to business and industrial life and one that arithmetic can make as well. It saves more time in the case of drills on reducing fractions to higher and lower terms alone than was required to learn its meaning and use. To rewrite a quantitative problem as an equation and then make the easy selection of the necessary technique to solve the equation is one of the most universally useful intellectual devices known to man. The words 'equals', 'equal', 'is', 'are', 'makes', 'make', 'gives', 'give', and their rarer equivalents should therefore early give way on many occasions to the '=' which so far surpasses them in ultimate convenience and simplicity.

(5) *Addition and Subtraction Facts in the Case of Fractions.*—In the case of adding and subtracting fractions, certain specific bonds—between the situation of halves and thirds to be added and the responses of thinking of the numbers as equal to so many sixths, between the situation thirds and fourths to be added and thinking of them as so many twelfths, between fourths and eighths to be added and thinking of them as eighths, and the like—should be formed separately. The general rule of thinking of fractions as their equivalent with some convenient denominator should come as an organization and extension of such special habits, not as an edict from the text-book or teacher.

(6) *Fractional Equivalents.*—Efficiency requires that in the end the much used reductions should be firmly connected with the situations where they are needed. They may as well, therefore, be so connected from the beginning with the gain of making the general process far easier for the dull pupils to master. We shall see later that, for all save the very gifted pupils, the economical way to get an understanding of arithmetical principles is not, usually, to learn a rule and then apply it, but to perform instructive operations and, in the course of performing them, to get insight into the principles.

(7) *Protective Habits in Multiplying and Dividing with Fractions.*—In multiplying and dividing with fractions special bonds should be formed to counteract the now harmful influence of the

'multiply = get a larger number' and 'divide = get a smaller number' bonds which all work with integers has been reinforcing.

For example, at the beginning of the systematic work with multiplication by a fraction, let the following be printed clearly at the top of every relevant page of the text-book and displayed on tbc blackboard.

When you multiply a number by anything more than 1 the result is larger than the number.

When you multiply a number by 1 the result is the same as the number.

When you multiply a number by anything less than 1 the result is smaller than the number.

Let the pupils establish the new habit by many such exercises as:—

$18 \times 4 = \ldots$	$9 \times 2 = \ldots$
$4 \times 4 = \ldots$	$6 \times 2 = \ldots$
$2 \times 4 = \ldots$	$3 \times 2 = \ldots$
$1 \times 4 = \ldots$	$1 \times 2 = \ldots$
$\frac{1}{2} \times 4 = \ldots$	$\frac{1}{3} \times 2 = \ldots$
$\frac{1}{4} \times 4 = \ldots$	$\frac{1}{6} \times 2 = \ldots$
$\frac{1}{8} \times 4 = \ldots$	$\frac{1}{9} \times 2 = \ldots$

In the case of division by a fraction the old harmful habit should be counteracted and refined by similar rules and exercises as follows:

When you divide a number by anything more than 1 the result is smaller

When you divide a number by 1 the result is the same as the number.

When you divide a number by anything less than 1 the result is larger than the number.

State the missing numbers:—

$8 = \ldots 4s$	$12 = \ldots 6s$	$9 = \ldots 9s$
$8 = \ldots 2s$	$12 = \ldots 4s$	$9 = \ldots 3s$
$8 = \ldots 1s$	$12 = \ldots 3s$	$9 = \ldots 1s$
$8 = \ldots \frac{1}{2}s$	$12 = \ldots 2s$	$9 = \ldots \frac{1}{3}s$
$8 = \ldots \frac{1}{4}s$	$12 = \ldots 1s$	$9 = \ldots \frac{1}{9}s$
$8 = \ldots \frac{1}{8}s$	$12 = \ldots \frac{1}{2}s$	
	$12 = \ldots \frac{1}{3}s$	
	$12 = \ldots \frac{1}{4}s$	

$16 \div 16 =$	$9 \div 9 =$	$10 \div 10 =$	$12 \div 6 =$
$16 \div 8 =$	$9 \div 3 =$	$10 \div 5 =$	$12 \div 4 =$
$16 \div 4 =$	$9 \div 1 =$	$10 \div 1 =$	$12 \div 3 =$
$16 \div 2 =$	$9 \div \frac{1}{3} =$	$10 \div \frac{1}{5} =$	$12 \div 2 =$
$16 \div 1 =$	$9 \div \frac{1}{9} =$	$10 \div \frac{1}{10} =$	$12 \div 1 =$
$16 \div \frac{1}{2} =$			$12 \div \frac{1}{2} =$
$16 \div \frac{1}{4} =$			$12 \div \frac{1}{3} =$
$16 \div \frac{1}{8} =$			$12 \div \frac{1}{4} =$
			$12 \div \frac{1}{6} =$

(8) *Per Cent of Means 'Hundredths Times'.*—In the case of percentage a series of bonds like the following should be formed:

5	per cent of	= .05 times
20	" " "	= .20 "
6	" " "	= .06 "
25%	"	= .25 ×
12%	"	= .12 ×
3%	"	= .03 ×

Four five-minute drills on such connections between '*X* per cent of' and 'its decimal equivalent times' are worth an hour's study of verbal definitions of the meaning of per cent as per hundred or the like. The only use of the study of such definitions is to facilitate the later formation of the bonds, and, with all save the brighter pupils, the bonds are more needed for an understanding of the definitions than the definitions are needed for the formation of the bonds.

(9) *Habits of Verifying Results.*—Bonds should early be formed between certain manipulations of numbers and certain means of checking, or verifying the correctness of, the manipulation in question. The additions to $9 + 9$ and the subtractions to $18 - 9$ should be verified by objective addition and subtraction and counting until the pupil has sure command; the multiplications to 9×9 should be verified by objective multiplication and counting of the result (in piles of tens and a pile of ones) eight or ten times,* and by addition eight or ten times;* the divisions to $81 \div 9$ should be verified by multiplication and occasionally object-

*Eight or ten times *in all*, not eight or ten times for each fact of the tables.

ively until the pupil has sure command: column addition should be checked by adding the columns separately and adding the sums so obtained and by making two shorter tasks of the given task and adding the two sums; 'short' multiplication should be verified eight or ten times by addition; 'long' multiplication should be checked by reversing multiplier and multiplicand and in other ways; 'short' and 'long' division should be verified by multiplication.

These habits of testing an obtained result are of threefold value. They enable the pupil to find his own errors, and to maintain a standard of accuracy by himself. They give him a sense of the relations of the processes and the reasons why the right ways of adding, subtracting, multiplying and dividing are right, such as only the very bright pupils can get from verbal explanations. They put his acquisition of a certain power, say multiplication, to a real and intelligible use, in checking the results of his practice of a new power, and so instill a respect for arithmetical power and skill in general. The time spent in such verification produces these results at little cost; for the practice in adding to verify multiplications, in multiplying to verify divisions, and the like is nearly as good for general drill and review of the addition and multiplication themselves as practice devised for that special purpose.

Early work in adding, subtracting, and reducing fractions should be verified by objective aids in the shape of lines and areas divided in suitable fractional parts. Early work with decimal fractions should be verified by the use of the equivalent common fractions for .25, .75, .125, .375, and the like. Multiplication and division with fractions both common and decimal should in the early stages be verified by objective aids. The placing of the decimal point in multiplication and division with decimal fractions should be verified by such exercises as:—

$$\begin{array}{r} 20 \\ 1.23 \;\big|\overline{\;24.60.} \\ 246 \\ \hline \end{array}$$

It cannot be 200; for 200×1.23 is much more than 24.6. It cannot be 2; for 2×1.23 is much less than 24.6.

The establishment of habits of verifying results and their use is very greatly needed. The percentage of wrong answers in arithmetical work in schools is now so high that the pupils are often being practiced in error. In many cases they can feel no genuine

and effective confidence in the processes, since their own use of the processes brings wrong answers as often as right. In solving problems they often cannot decide whether they have done the right thing or the wrong, since even if they have done the right thing, they may have done it inaccurately. A wrong answer to a problem is therefore far often ambiguous and uninstructive to them.

These illustrations of the last few pages are samples of the procedures recommended by a consideration of all the bonds that one might form and of the contribution that each would make toward the abilities that the study of arithmetic should develop and improve. It is by doing more or less at haphazard what a psychology teaches us to do deliberately and systematically in this respect that many of the past advances in the teaching of arithmetic have been made.

A CRITICAL STUDY OF THE CONCEPT OF SILENT READING ABILITY

L. W. PRESSEY and S. L. PRESSEY.
University of Indiana.

I. The Misleading Character of Many Test "Labels"

As has been well pointed out recently[1] by Thorndike and Courtis, the name which is applied to a test may often be positively misleading, in its implications as to what functions that test measures. A test consisting of problems in addition might seem unmistakably a test of ability to add; but careful investigation may discover that comparative standing on this test is determined almost wholly by speed in writing the answers, or knack in working ahead of one's pencil. And the determination of true ability in addition from any addition test is found to be by no means a simple matter.

If such difficulties appear in the case of such a very concrete and specific subject as arithmetic it is surely reasonable to infer even greater likelihood of error in efforts at measurement in such subjects as literature, composition, silent reading, and so on. Particularly (the writers have felt) was the situation confused and ill-defined with regard to tests in silent reading, and to "silent reading ability." There are here no very specific habits, as in measurement of ability in the fundamentals of arithmetic, which make very specific materials for investigation. Nor are there—as "silent reading ability" is usually conceived—any informational elements, as in history or geography, about which investigation may be centered. Under such circumstances, the formulation of tests which would in truth measure "silent reading ability" would seem a peculiarly difficult task. It is surely easy enough to devise a test in which the element superficially most prominent is ability to read with understanding. We may then, with a certain nonchalance, label such a test a "test of silent reading ability." But to a much greater extent than with addition tests is there danger that our "label" may be inaccurate and even harmfully misleading. Critical research with regard to such tests would seem, then, all the more needful.

---

[1]Thorndike E. L. and Courtis S. A., "Correction Formulae for Addition Tests," *Teachers' College Record*, 1920, 21, 1-24.

It is such a critical study that the present paper reports. Previous research with silent reading tests had led the writers to feel that the situation, with regard to silent reading, needed to be analyzed. Particularly, they felt that ability in assimilative reading was conditioned much more by the nature of the material read than is ordinarily supposed. The problem of the present paper may, then, be put as follows: *is either the form or the content of the matter read an important conditioning factor in silent reading?* If the nature of the material read should appear to be a dominant feature in the situation, it would seem that our tests of general silent reading ability were badly mislabeled.

II. Materials and Results

The materials of this study were obtained with four reading "scales" each composed of a definite type of reading material. The items for three of them were taken from the Monroe Reading Scales and the "Illinois Examination"[2]; some of the items in the fourth "scale" also came from these two scales. Each "scale" had twelve items. One "scale" was made up entirely of poetry, one of scientific passages and two of ordinary reading matter, stories, and the like. The two "general" scales were made by simply taking alternate items from the Illinois Examination for the 3rd, 4th and 5th grades, the same examination for the 6th, 7th and 8th grades and from the two forms of Monroe's reading tests for the 6th, 7th and 8th grades. The items involving poetry were omitted. These two scales were roughly equal in difficulty, and of a very similar type of reading matter. An equivalent "scale" was then made up using the poetical items. Very few "scientific" items could be found in the scales, but those that appear were used, others being made to fill out the necessary number of items for a fourth "scale". The distribution of scores and the medians for these four scales were practically the same; the medians of the four are,—6.1, 6.5, 6.6, 6.4. It seems reasonable to suppose, then, that these scales are of about equal difficulty. Samples from these four "scales" appear below:

From the poetical scale

O suns and skies and clouds of June,
  And flowers of June together,
You cannot rival for one hour
  October's bright blue weather.

[2]See Journal of Educational Research, October, 1920. The Illinois Examination contains a revision of the original Monroe Scale.

Which month does this stanza say is the more pleasant?
April September May June October

From the first "general" scale

The caravan, stretched out upon the desert, was very picturesque; in motion, however, it was like a lazy serpent. By and by its stubborn dragging became intolerably irksome to Balthasar, patient as he was.

Place a line under the word which tells in what respect the caravan resembled a serpent.
temper color length motion size

From the second "general" scale

It was the garden-land of Antioch. Even the hedges, besides the lure of the shade, offered passers-by sweet promises of wine and clusters of purple grapes. Over melon patches and through apricot and fig groves and groves of oranges and lime, the white-washed houses of the farmers were seen.

What kind of a land was this? Draw a line under the correct answer.
barren hilly productive infertile desert

From the "scientific" scale

The tighter a wire is stretched, the higher the tone produced will be when the wire is struck. Five wires are stretched with weights on the ends of them. One weight is 100 lbs., one is 75 lbs., one is 25 lbs., one is 20 lbs. and one is 15 lbs.

Underline the number of lbs. in the weight that will cause the highest tone.
75 lbs. 100 lbs. 25 lbs. 20 lbs. 15 lbs.

These four "scales" were then given to a large seventh grade,[3] giving a total of 112 cases. The "poetical" scale was given first, followed by the first "general" scale; then came the "scientific"

---

[3]The writers wish to express their obligations to Superintendent E. W. Montgomery of Bedford, Indiana, for his kindness and co-operation in this research.

scale and the second "general" scale. By this arrangement, practice with the tests, fatigue and so on were so distributed that they could not act as a constant error in the results. The tests were scored simply on the basis of the total number of items right with each type of reading matter, previous research having suggested that "rate" score was negligible as a measure of silent reading ability, independent of comprehension score.[4] All possible correlations between the four "scales" were then found. The coefficients appear below:[5]

Correlation of "general 1" with "general 2".	.85
Correlation of "general 1" and "poetry"	.38
Correlation of "general 1" and "scientific".	.35
Correlation of "general 2" and "poetry".	.31
Correlation of "general 2" and "scientific".	.49
Correlation of "scientific" and "poetry".	.56

The exceedingly high correlation between the two "general" scales is first to be noted. This coefficient gives a measure of the "reliability" or "consistency" of the scores yielded by the scale. It seems reasonable to believe that two "poetry" or two "scientific" scales would yield scores that correlated to somewhat the same extent. Similar types of reading matter would probably correlate highly. The writers happen to know that this coefficient of reliability is somewhat higher than is usually found with the Monroe scales,—the coefficients usually run from .65 to .75. It might be suggested that it is the presence of items involving poetry or scientific reading that lowers these correlations.

The striking fact, however, is the contrast between the very high correlation above noted and the remaining coefficients. Strictly, the most reliable coefficients are those between "general 1" and the "scientific", and between "general 1" and "poetry", as contrasted

---

[4]Pressey, S. L. and L. W., "The Relative Value of Rate and Comprehension Scores on Monroe's Silent Reading Test." School and Society, June, 1920. Since that paper was published, a second study has been made confirming the finding as to the negligible nature of rate score. *It also indicates that Monroe's weighting of his comprehension scores adds practically nothing to a crude score as a measure of reading ability.* As a result of these two studies, crude scores,—that is, comprehension scores—have been used throughout this study with no attempt at any more elaborate treatment.

[5]It should be added, in this connection, that a preliminary trial of the four scales in a consolidated country school, testing grades 5, 6, 7 and 8, gave practically identical results. Only 67 children were included in the correlation, and the group was heterogeneous, so the correlations are not as reliable as those given in the paper. But they make very pretty confirmatory evidence. The correlation between the two "general" scales was .83, and none of the others were above .49.

with the correlations between "general 1" and "general 2"; the most significant contrast appears here. That is, the "poetry" scale immediately preceded "general 1" while the "scientific" passages immediately followed it. Adventitious circumstances such as fatigue should, then, influence these correlations less than those between scales further apart in the series. The correlation between "general 1" and "general 2" is, thus, a correlation between scales somewhat apart in the series; this correlation is thus probably slightly lower than it should be, while the correlations between "general 1" and the other two scales express to the full the relation between ability to read differing types of subject matter.

The results are surely striking. The findings need confirmation of course. Study of reading ability in its relation to vocabulary in a particular field, study of ability in assimilative reading interesting and not interesting to a given child, studies in "assimilative listening" as compared with ability in assimilative reading,—there are a dozen and one bits of investigation which at once suggest themselves as likely to throw further light on the problem. Meanwhile, certain implications of the data just presented are worth considering.

III. Need for a Re-formulation of the Silent Reading Problem

It appears, then, that ability in silent reading depends very largely upon the nature of the passage read; a good reader in one type of subject matter may very likely be a poor reader with other material. And the usual silent reading scale may be considered to measure—certainly not silent reading ability in general, since there seems to be little evidence of any general factor of outstanding importance. Nor do the standard scales specialize on one type or other of subject matter; it might possibly be questioned whether they measured much more than ability in dealing with such catechistic series of questions. The writers would hardly go to such an extreme conclusion. But they do feel that the results suggest a distinct revision of present concepts with regard to silent reading ability and its measurement. As the writers see it, "silent reading ability" involves four distinct factors.

There must be (1) freedom from oral reading habits. It is evidently of great importance to achieve such freedom. This factor is a negative one, however, and is not to be considered correlative

with a positive ability. Tests for detecting the presence of oral reading habits would hardly take the form of the usual reading test. Further, after freedom from such handicaps has been won, the factor should be negligible; one is then above the "threshold" in silent reading ability. Silent reading ability, defined as assimilative reading independent of oral reading habits, needs no more to be stressed than phonetics in the sixth grade.

A second factor (2) in efficient assimilative reading is a large reading vocabulary and a background of information. Evidently a reading vocabulary is the factor most readily acquired by reading. Measurement of vocabulary should, the writers feel, be a regular practice in appraising the work of a teacher or school. By a curious and unfortunate circumstance, vocabulary tests have received most attention as measures of general intelligence. The acquirement of an adequate reading vocabulary, presumably by a large amount of required reading of a type in which the pupils are interested, would seem to the writers a matter to which the schools should give decided attention. In fact, they feel that this is the neglected element in the silent reading situation.[6]

A third factor (3) of prime importance is, of course, the development of interest to motivate reading of the right kind. Once there is such interest, a pupil will usually shake himself free of oral reading habits in his eagerness to cover the subject matter in which he is absorbed. And if there is interest, the acquirement of a reading vocabulary will come of itself. The writers have a distinct feeling that silent reading is now a problem in the public schools largely because the reading matter required by the schools is not such as to interest the pupils. If there had been interest, the two previously mentioned factors would, in most instances, take care of themselves.

The fourth factor (4) which conditions efficiency in assimilative reading may be considered the development of habits of attention and application. But these factors are factors in silent reading only incidentally and training for this factor in silent reading is, of course, training in methods of study.

What, then, about the present tests in silent reading? It is the

[6]It might be added that the writers have included such a test in a scale, designed to measure achievement in the eighth grade. A full statement with regard to this scale will be published shortly. The material reported in the present paper was, in fact, gathered in connection with the construction of this scale.

writers' guess that scales of the type of the Kansas Test and the Monroe Test are really by far the best examples so far of tests of attention,—which the devisors have stumbled upon without knowing it. They are good tests, but they have the wrong label. They are useful in investigating the efficiency of study habits. They may be used in hunting out those who are not yet free from oral reading habits—though instruments more explicitly devised for this one purpose should be much more efficient. For investigation of real ability in assimilative reading the writers would suggest: (a) a preliminary instrument for detecting oral reading habits; and (b) a test of vocabulary,—a classified test covering specific fields of information would be most valuable. These two tests might well be supplemented by (c) any one of the present standard reading tests. It should be clearly understood, however, that these last tests are used primarily to investigate habits of attention, not any process peculiar to the reading process alone.

These are the writers' conclusions on the basis of their data. Their findings need confirming. Meanwhile, tentative though these conclusions are, they have an interesting bearing on the general problem. They also suggest certain obvious shifts in present methods of teaching reading, and in interpreting the results with reading tests which are of no little interest.

A COMBINED MENTAL-EDUCATIONAL SURVEY *

RUDOLF PINTNER and HELEN MARSHALL,
Ohio State University.

During the last twenty years there have been two relatively independent movements within the field of psychology which have had to do with measurement, namely, the measurement of mental ability and the measurement of educational attainment. Although overlapping in many respects, it is only recently that any attempt has been made to combine the methods and results of these two lines of investigation in order to answer the question as to how much educational attainment we have a right to expect from a child of a given mental ability.

Children vary greatly in mental ability and our mental tests give us a fairly accurate measure of their native ability. These tests are measures of what children are capable of achieving rather than of what they have achieved. They are measures of the degree of modifiability of the individual. They are measures of the raw material with which the school and the teacher have to work.

Educational scales give us a measure of the amount of attainment in any specific school subject. They measure the extent to which the school has been successful in modifying its pupils. They give us the actual facts as to how much has been attained by each pupil.

These two measures, as they stand, are both extremely valuable. They increase tremendously in significance, however, as soon as we use the two measures in combination, and evaluate the one in the light of the other. If we make such an evaluation, we can expect answers to questions of this kind: How much educational attainment can I expect of this child who possesses so much general intelligence? If a child is able to do so much on our intelligence tests, how much of the ordinary school subjects ought he to know if given the common school instruction? Is this child of high intelligence accomplishing school work commensurate with his mental ability? Or is he working at a lower level, because he is in a grade that does not call for as much as he really could accomplish? Is this

---

*The two test blanks described in this article, namely, the Non-Language Mental Test and the Educational Survey Test, together with stencils for scoring and Manual of Directions giving tables of norms, may all be obtained from the College Book Store, Columbus, O.

school doing work worthy of the native ability of its pupil-material? Here is a school doing good work as judged by the average school in the system and also by the norms of standard educational tests, but is it really working up to capacity, or is it able to achieve such results easily and without effort because of the superior intelligence of its pupil-material?

We can see from the type of question here suggested, that we are raising fundamental questions, which the intelligence test alone or the educational test alone is unable to answer. We are attempting to use every bit of native ability of the pupil and by so doing, eliminate the waste of intelligence that is going on in our schools today. This wasted intelligence is much greater than most of us have realized. Ever since the introduction of mental and educational tests, we have all been aware of this wastage, but we have not been able to measure it accurately. The right combination of mental and educational tests will enable us to do so. It will show us just where this wastage exists, and its amount. It will show us just where—upon what school or class or pupil—pressure must be brought to bear, in order to eliminate the waste and make the school, class or pupil work up to the extent of the native capacity.

As a matter of fact we shall see, what we have long suspected and what we have only recently consciously realized, that the greatest amount of waste exists among the brighter pupils in a class or among the better schools in a school system. It is, as a rule, the more intelligent pupils that are working below capacity, even although they are keeping well up to the average level of the group. We have been pushing and cramming the duller children, while the brighter ones have been allowed to loaf. The bright child is the most retarded child in our schools. The dull child is the most accelerated. The bright child is the laziest child, and the dull child is the most industrious. We are now ready by means of combined mental and educational tests to equalize the pressure, and stimulate equally both the bright and the dull so that they may work up to their respective capacities. If we can accomplish this we shall have happier and better children. Lack of adequate stimulus undoubtedly leads to much disciplinary trouble in our schools. Habits of mental laziness acquired in school often persist through life, and there are undoubtedly many adults at the present time, who have splendid

native ability and who do not know it, because the school has taught them to be satisfied with a mediocre type of accomplishment.

The Tests.—The use of mental and educational tests in combination for the purposes outlined above was suggested by one of the writers* in 1918, arising naturally out of the plea for a wider use of group tests for survey purposes. Attempts to use the standard educational tests, as they existed at that time, for this combined mental-educational measure were found to be impossible, owing to the fact that practically no age norms for the tests had been published. Being tests of school achievement most authors had been content with grade norms only. In addition to this, most of the educational tests required a long time to give and several hours would be necessary in order to cover the chief subjects of the elementary school curriculum. It was, therefore, decided to devise a short educational survey test covering the chief elementary school subjects. This was composed of eight tests made up from standard educational tests and requiring about 30 minutes for its application. The tests, out of which this short composite educational test was composed, are (1) Thorndike's Vocabulary, (2) Woody's Arithmetic, (3) Kelly's Reading, (4) Thorndike's Reading, (5) Trabue's Language Completion, (6) Starch's Punctuation and Grammar, (7) Hahn and Lackey's Geography, (8) Van Waganen's History. The composition of this test and its relation to the longer standard tests have been discussed elsewhere.**

For a group mental test to go along with this educational survey test, it was decided to construct a test which would have as wide an application to as many different types of children as possible and which would also be far removed from ordinary school work. For this purpose a non-language group test was devised. It has been described in detail elsewhere.*** This test has been found useful for all school children from grade II upwards. It can be given with equal fairness to English-speaking as well as to non-English speaking or deaf children. It does not overlap at all with the type of thing that is taught in school. By avoiding reading and language tests, we can gauge more accurately the intelligence of children who have been deprived of schooling or whose schooling has been

*Pintner, R. The Mental Survey, Appletons, 1918.

**Pintner, R., and Fitzgerald, F. An Educational Survey Test. Journal of Educational Psychology, Vol. XI. No. 4, April, 1920, pp. 207-223.

***Pintner, R. A. Non-Language Group Intelligence Test. J. of Applied Psych. Vol. III. Sept., 1919. pp. 199-214.

very irregular or who have special disabilities in reading and, when we have done so, we can estimate what they ought to be able to accomplish in school. We shall likewise avoid the error of overrating the intelligence of the bookish child, who has had undue stimulation at home from the reading and language side.

The Standardization.—Standards for such a combined measure must be based upon children who have taken both tests. This is absolutely necessary, because of the wide variability in mentality and in educational accomplishment found among different groups of children. Only when we are sure that our standards are based upon a large and unselected sampling of school children, can we evaluate educational tests in terms of mental tests, and it is very seldom that we can be sure of any such thing. It was, therefore, decided to make a standardization based solely upon those cases that had been given both tests.

Altogether 4303 cases were tested on both tests, ranging in age from 6 to adult. There were from 300 to 600 cases at the intermediate ages and fewer at the upper and lower ends of the standardization. The distribution of the scores on both tests seemed to indicate an approach to a normal distribution when examined age by age. There was likewise a reasonable increase in both tests from one age to the other.

The Method of Rating.—The percentile rating is undoubtedly the most accurate and useful rating for mental testing purposes. In this case, however, we are dealing with two tests and we are primarily concerned with the difference between the ratings in these two tests, in order that we may be able to state by how much a child in his educational accomplishment falls below his mental ability or vice versa. If we have his percentile ratings on the mental and on the educational tests, the difference between these two percentile ratings will be the type of measure we want, except that it will be very inaccurate. In as much as our distribution approaches the normal, a difference of five points near the median will be of much less significance than the same difference at either end of the distribution. For example, if one child makes percentiles of 50 and 45 on the mental and educational tests, respectively, and another 85 and 80, both will show a difference of 5 points. Both are 5 points lower educationally than they are mentally. But the significance of this

difference of 5 points is very different in the two cases. In the center of our distribution there is little difference in actual ability between the 45th and the 50th case; but there is a great difference in ability between the 80th and the 85th case. It is obvious, therefore, that the percentile method, in spite of its many advantages, cannot be used for our purpose.

The method adopted was, therefore, to assume a normal distribution of the scores and convert the percentage frequency of any score into terms of the mean square deviation or sigma. By this means all scores are evaluated in terms of sigma and the difference between these sigma values is comparable at any part of the surface of distribution. This procedure was carried out for both tests for all ages. It would, however, be rather inconvenient and decidely puzzling to the average teacher to be asked to think in terms of plus and minus sigma, and particularly so when we are handling the results of two tests expressed in sigma values and when we require the difference between these two sigma values to stand for the excess or deficiency in the mental test over the educational.

To avoid this complexity the sigma values were turned into a scale of so-called index numbers running from 0 to 100, zero being placed at the lower end of the distribution and 100 at the other, and 50 corresponding to the zero point of the normal probability surface. If now we multiply the sigma values by a constant and add or subtract the product from 50 we obtain a new scale of numbers, which we have called index values. These values are all positive whole numbers and the teacher can readily understand their significance. It is easy for her to grasp the meaning of a mental index of 50 and an educational index of 47 and a difference between the two of minus three, which can be interpreted as meaning that the particular child is three points lower in educational attainment than he should be as compared with his native ability. And furthermore that this child with a difference of —3 is not nearly so poorly adjusted as another child who shows a difference of —12 after subtracting his educational index of 70 from his mental index of 82, even although the mental ability and educational accomplishment of the latter child are far above those of the first case. In other words, the teacher, after scoring the tests, converts the scores by means of a table into index values and the significance of the index values is explained as follows:

Index Values	Mental	Educational	Per cent of cases
0 to 19	Dull	Very Poor	2
20 to 39	Backward	Poor	23
40 to 59	Normal	Average	50
60 to 79	Bright	Good	23
80 to 100	Very Bright	Very Good	2

That is to say, all index values between 0 and 19 on the mental test show dull mentality and on the educational test indicate very poor educational accomplishment, and fall among the lowest 2 per cent of the cases tested; similarly index values from 20 to 39 indicate backward mentality or poor educational accomplishment and fall among the lower 23 per cent of the cases, and so on. The examiner soon becomes familiar with this table, because it is very easy to remember. The first 20 index numbers refer to the lowest group, the next 20 to the lower group, the next 20 to the average or intermediate group, the next 20 to the higher group and the top 20 to the highest group.

The Three Ratings.—Having given the two tests and having calculated the indices and the difference, the examiner now has three valuable ratings:

1) The Mental Index, which is a measure of the native ability of the child.
2) The Educational Index, which is a measure of the school attainment of the child.
3) The Difference, which expresses the relationship between the two previous ratings.

The first two ratings have already been interpreted and the meaning of an educational or mental rating is now thoroughly familiar to all teachers. The new rating that we have introduced here is what we have called the Difference, i. e., the difference of a child's standing in the educational and mental tests, or the difference between his native capacity and his actual accomplishment. Such a rating we believe to be the most important contribution that the science of mental measurement can offer to the educator. This Difference is arrived at by subtracting the mental from the educational index. If the mental index is higher the Difference is marked by a minus.

If the educational index is higher the Difference is marked by a plus. A minus difference means that the child is doing less educational work than he has the ability to accomplish. A plus difference means that he is doing more educationally than has usually been accomplished by children of like mentality. Notice the significance of the plus difference. Because our educational test norms are based upon what is actually now being accomplished by a fair sampling of school children, we have no measure of what ought to be accomplished under ideal conditions where each child is working up to the limit of his capacity. It is useless to attempt to set up any such ideal standard. We, therefore, find many cases with a plus difference and this means that these cases are children who are accomplishing in school work more than is usually accomplished by children of like intelligence.

The Significance of the Difference Ratings.—In order to interpret these Differences, a frequency distribution of the Differences of 4303 cases was made. We may conveniently divide this frequency distribution into a five-fold grouping, as follows:

Differences	Percentage of Cases
—24 and below	2.5
—23 to — 9	19.0
— 8 to + 8	53.5
+ 9 to +23	22.0
+24 and above	3.0

In other words, Differences ranging from minus 8 through zero to plus 8 are found in about 50 per cent of the cases and are, therefore, not significant. This is the condition found in the mass of school children at the present time, and without assuming this to be an ideal condition, or without assuming that this group is actually working up to its mental capacity, we must, nevertheless, accept the facts and from this basis of fact work toward a saving of the waste of intelligence that is apparent in the cases showing a difference of minus 9 or below.

The 21.5 per cent having Differences of minus 9 or below are obviously marked cases of wasted intelligence. They have ability to do much better and for some reason or other this ability is not being utilized. It is the business of the school to find out the reasons and

to stimulate them further. There may be many reasons, but, one reason, as we shall see in our second paper, is the very obvious one of misplacement in grade. Many of them are in a lower grade than they ought to be. Those showing Differences of minus 24 and below make up about 2.5 per cent of our cases and are, of course, the extreme examples of mal-adjustment. Prompt study of these cases is needed. They show the extreme of wasted intelligence.

Those showing Differences of plus 9 and above are children who are accomplishing more than is usually expected of children of their mentality. Again there are many reasons for this. One common reason is the pressure brought to bear on the slower child in order to make him keep pace with the grade suited to his chronological age, and the enormous amount of extra attention and especially skillful teaching that is lavished on this child. For these particular cases that show Differences of plus 9 and above, we believe that this emphasis is wrong. It is false economy at the present time when so much more intelligence is being allowed to run to waste.

The significance of the Difference ratings of schools will be somewhat different than that for individuals, because the Difference of a school is derived from the median mental and educational indices of all the children in the school. This decreases the spread of difference values. In all, 56 different schools were tested and the middle 50 per cent show differences from plus 3 to minus 3. For schools or classes a difference of medians above or below 3 becomes significant. All such schools having difference values greater than minus 3 show that educational attainment is too far below mental capacity and improvement should be expected.

The Grade Rating.—Our standardization of the educational tests has not been restricted to an age standardization merely. We have also prepared a grade standardization. This follows the scheme of the previous standardization, so that we may convert any educational score into an educational index according to grade. In doing this we compare the score obtained by a specific child with the other scores obtained by children of like grade and read off the index from the table. The indices have the same significance as before, namely, 0-19 meaning very poor work for that grade; 20-39 poor work; 40-59, average work; 60-79, good work; and 80-100, very good work.

In this way a teacher can compare children or classes with refer-

ence to their grade attainment. As we shall see, this educational grade index becomes important in deciding which children should be skipped or accelerated.

Mental or Educational Age.—If we use the scheme of indices as explained above there is, of course, no need to convert our results into mental or educational ages. Those who wish to do this, however, can readily do so by using the 50 index line in the index table. This gives the median scores for each age. Any particular score may thus be converted into a mental age. For example, if a child makes a score of 290 and if this is the median for age 11, the mental age will be 11. We may interpolate between ages and arrive at fractions of any age. In exactly the same way we may compute an educational age from the 50 index or percentile line of the table of educational scores.

Having these mental and educational ages, we can then in the usual manner calculate an intelligent quotient and an educational quotient, respectively. We may then use this I. Q and E. Q. for any purpose we wish. The I. Q. divided by the E. Q. will give us a ratio, the significance of which will be similar to the Difference we have described above, only, of course, the values will cluster around 1.00, ranging upwards and downwards, instead of clustering around 0 with plus and minus quantities. The conversion of scores into mental and educational ages followed by the calculation of I. Q. and E. Q. and the ratio between them, involves much more work than the conversion of scores into indices and the simple subtraction of these indices. No advantage is gained by having ages and quotients, and we recommend the use of indices and differences on our tests.

The Reliability of the Tests.—One measure of the reliability of a test is the amount of correspondence between the rating of the same children tested by the same test at different times. The degree of reliability will be influenced by the homogeneity of the mentality of the group and by the time interval between the two testings. The less homogeneity the higher will be the correlation; the longer the time interval the lower the correlation.

The mental tests were given to a group of children in 1918 shortly after the tests had been devised. The same group were given the tests again in 1920. This group consists of 46 children ranging in chronological age from 10 to 14 in 1920, in mental indices from 28

to 81. Both the 1918 and 1920 tests were rated in accordance with the final standardization. What we are interested in knowing is to what extent the 1918 ratings are valid after two years have elapsed. Has a lapse of two years justified the ratings given in the first test? The coefficient of correlation between the two tests will denote the degree of correspondence in the ratings of the two testings. The correlation between the mental indices gives a coefficient of r =.72. In other words, the index that a child receives on the mental test is not likely to change much even after the lapse of two years.

If we rank the 46 cases according to their mental indices we find that the correspondence between them for the two testings is expressed by a rank coefficient of correlation of .77. In other words, we may say that on the whole a child does not vary greatly in rank after a lapse of two years. Some individual cases do show a great change in rank, while others maintain the same ranks, but on the whole the change is not great.

Of the 46 cases, 26 showed a gain in their mental index from 1918 to 1920, and the average gain for these 26 was 6.5 points; 18 showed a loss in mental index, and the average loss for these 18 was 5.8 points; while 2 had exactly the same index in both years. The greatest change in mental index was 20 points. This was made by a child who had a mental index of 27 in 1918 and 47 in 1920.

The correlation between the crude scores obtained on the tests will not be as great as that between the mental indices, and the longer the interval between any two testings the lower will the correlation for any given group become. This is due to the fact that children of different mental ability are progressing at different rates and the increase in crude score will vary according to the mental ability of the child. The Pearson Product Moment correlation for the crude scores for the 1918 and 1920 testings is 50 and the rank correlation is 58.

We may sum up this discussion of the reliability of the tests by saying that they show a remarkable degree of correspondence in the ratings of the children after a long period of time. This must not be interpreted as meaning that the index of any one child will remain constant over a long period. There are too many extraneous and accidental factors at work in group testing to warrant any such statement. Under good conditions, however, the chances seem to be

greatly in favor of rather close correspondence even after as long an interval as two years.

Agreement with Binet Ratings.—For 300 children, ranging from the 3rd to the 6th grade, a correlation of .47 was obtained between the mental index on the mental test and the I. Q. on the Stanford Binet. All these cases were tested during the same semester. This correlation between a 30 minute group examination and the thoroughgoing individual examination of the Stanford Binet is very satisfactory. We must bear in mind also the radical difference between the two tests, not only as to method of giving, e. g. group versus individual, but also as to content, e. g. non-language absolutely versus predominantly language responses.

The literary and scholastic nature of the Binet test seems reflected in the higher correlation between the indices obtained by the same group of children on the Educational Survey Test and the I. Q. on the Binet. In this case the coefficient is .52.

The correlation between absolute achievement on both tests as expressed by crude score on the non-language group test and mental age on the Binet is .80, which is very high and seems to show that the group test is discriminating well between different degrees of mental ability.

The correlation of the same group of cases between mental and educational indices is .62, which shows that on the whole educational attainment tends to go with mental ability, but that there is plenty of room for improvement, and therefore plenty of work for the psychologist to do in helping to adjust the child to the work he requires.

The correlation of the same group of cases between the Differences and the I. Q.s is .21. This shows as we shall demonstrate concretely in a later article, that we find children of all degrees of mental ability who are wasting their time in school and who are not working up to their real mental capacity.

Summary.—1) We have tried to show that the next logical step in psychological and educational measurement is the combination of mental and educational tests.

2) We have prepared two group tests, an educational and a mental, to measure school work and native ability respectively.

3) We have standardized these two tests and worked out a simple method for estimating the Difference between them.

4) This Difference is the most important value for school diagnosis, because it expresses the relationship between achievement and ability.

5) Untold waste in mental ability is brought to light by this method of study.

6) All intelligent school supervision will have to take account of the difference between ability and achievement and see to it that the educational achievement of each child is worthy of his ability.

DEPARTMENT FOR DISCUSSION OF RESEARCH PROBLEMS

Conducted by LAURA ZIRBES

This department has a two-fold function. It aims to serve research workers as well as educators, whose work brings them in close contact with children in the schools. It hopes to accomplish this service by suggesting research studies, which will meet well-defined school needs.

In order that this service may be real and effective, the co-operation of research workers and school people is desired. Correspondence with reference to the following questions will be considered in selecting topics for future discussions.

a. Which of the studies proposed would help you to solve a practical problem?

b. What topics might well be added to this list? Replies may be addressed to: Miss Laura Zirbes, 646 Park Ave., New York City.

A PROBLEM FOR CAREFUL RESEARCH:

How Does the Ability to Make Rapid Mental Adjustments Affect Test Scores and Pupil Achievements?

The problem stated in the title is one which was suggested by a number of class room situations met during the past three years. A careful inquiry should shed light on a question which has many definite bearings on teaching practice and educational outcomes.

The problem first took definite form in connection with the use of Courtis practice tests in arithmetic in nineteen seventeen. It has since become significant in a number of connections with various types of material. A brief discussion of such instances will lead to a definition of the scope of the problem and to restatements in terms of concrete situations.

***Instances from Arithmetic.*—The Courtis practice materials were being used in grades IV, V and VI. A study of the errors made revealed a great number of those listed by Mr. Courtis in the handbook which accompanies his practice material. It revealed also mistakes which did not fall under any of the headings for which practice is provided in the test series. These mistakes were most frequent on Test A—Lesson Thirteen. This indicated that there was a factor in Test A which was not so active in other lessons. The errors of a great number of pupils are so similar that it seemed possible that they were all due to one cause. The errors in question**

were of this sort. The subtraction examples which follow addition examples were worked as addition. The multiplication examples which follow the subtraction were worked either as addition or as subtraction. Tendencies to this sort of error were also noted in parts of examples, e. g. responses to 6×4 were 10 or 2. These errors were most frequent in the first few examples of each kind. There were some instances of pupils catching themselves in this type of error after a few examples had been worked. There were many cases in which the pupil could not see how the correct answer was possible until his mistake was pointed out.

A lesson was framed to give training in changing from one process to the other. This type of training is not provided in Courtis practice tests 1-12, inclusive, excepting as it occurs within the process of long division. The factor in question was played up in exaggerated form. Examples from Lessons 1-12, inclusive, were included. Time allowance was made to correspond with proportionate time allowed for examples in the respective lessons from which they were taken. Pupils took longer to work this lesson or did not finish in the prescribed time. The practice effects noted were increased accuracy and increased speed. Test A was then repeated with success at first trial. Courtis' suggestion is that pupils who fail to pass Test A need practice on Lessons 1-12. There was a strong indication that it is more economical to provide an extra practice lesson which would provide a type of training required by Lesson A but not supplied by Lessons 1-12.

The Woody arithmetic tests and Courtis Standard Research Test—Series B were given next. There were types of error in the Woody tests not found in the Courtis tests. Analysis showed this to be due to the grouping of similar forms in the Courtis test, and to the fact that in the Courtis Test the process could hardly be mistaken due to the character of the examples used and to the grouping. Each part of the Courtis Test, Series B could be worked with no reference to instructions as to the process required as each process is put in a separate test. This test then neglects absolutely a factor which is essential to success in Woody tests and in mixed process review lessons and tests. The Cleveland tests in arithmetic similarly ignore this factor, although the process cannot always be recognized by the form of the example. A mixed test was made composed of examples from the Cleveland spiral test. Both the

accuracy and speed scores were reduced by including different processes in one test.

These are the indications which suggest that it is necessary to train and test a skill factor which contributes to success, although it is not distinctly characterized as mathematical ability.

During the next year it became apparent that there was a similar situation and need for training due to the varieties of examples within each process. This made it necessary to supplement practice material provided by the Courtis tests in long division especially. For example, tests in which all the examples had zero difficulties were not difficult as soon as this fact was discovered. The same examples mixed with others in which there was no zero in the quotient showed errors which yielded to training with lessons which combined in one test, types of long division examples from various Courtis practice lessons. A type of example not provided in any of the tests was added because its entire omission gave the pupils a foreknowledge or check on accuracy not provided in other processes. The examples with remainders are the ones in question.

The detailed illustrations from the field of arithmetic make it possible to suggest more briefly the points at which a similar problem is presented in connection with other types of subject matter, and suggest investigation and experiment with may lead to solutions.

Illustrations from Reading.—A diagnostic study of the reading abilities of a fourth grade class was made during 1917-1918. Difficulties found were often due to a failure on the part of the child to set his mind for the particular reading task in hand, or to the persistence of one mental set through different types of reading. Concrete instances demonstrate the need of a type of training which will affect the ability to make mental adjustments rapidly and to see the necessity of change in mental set as it inheres in situations.

Pupils had very little silent reading experience. When the first measurement of oral and silent rate was made there was very little difference between the two scores for a great part of the class. Some of the differences between oral and silent reading were explained. Pupils were told that their lips might rest—silent reading is done by the eyes and the mind. Words are changed to pictures instead of sounds. Differences were illustrated, but no other in-

struction was given. Pupils were re-tested and there appeared variations which indicated that the mere consciousness of a need of change in mental set had an effect on rate of silent reading. Comprehension was checked so that the pupils were really reading at a faster rate silently, although the explanation made no mention of rate of reading. After this test, training designed to build up a *number* of silent reading "sets" rather than *one* was devised. Pupils were first given reading material which they were told to read with no further instructions as to the purpose of reading. They were then asked questions all of which could hardly be answered by a pupil who had maintained one "set" during the entire reading. The need of change was made inherent in the material, but the pupils did not see it.

The same material was repeated while pupils realized that their reading had not been effective. They were asked to check or underline parts of the material which they had not noticed before.

Next, a four part test involving change of set was devised. The first part was a very simple little story. The next was worded so that the sentences were not simple, although the words were not difficult. The next was an easy rhyme in which there was considerable repetition. The next was a problem. Each part was timed separately. There were some pupils whose rates in the four parts were nearly equal. The more effective readers judged by other comprehension and reproduction tests were those who varied their rate to suit the type of material. This indicated that there was a possible need of training in mental adjustment of this sort.

Analysis of the oral errors, which might have a bearing on comprehension, raised problems with regard to the possible defects of training which over-emphasizes one mental set. Some pupils disregarded word endings; others habitually read word by word without grouping into phrases; others skipped little words and went on regardless of meanings which were mutilated. Others were so intent on getting mental pictures and pleasurable experience that they paid no attention to precise words and often made interpretations which were not inherent in the text.

Brief Suggestions from Other Subjects.—It will be necessary to refer only briefly to the influence of mental set on spelling achievement. Pupils who spell well when spelling is uppermost in consciousness, become very inefficient spellers in circumstances which

demand the presence of other factors. Words learned for a given test or lesson can often not be recalled because the "set" in which they were acquired was not conducive to recall under other conditions.

Brief references to illustrations from geography illustrate the necessity for rapid mental adjustment in problem-solving wherever it occurs. For example, pupils were asked to outline the topics to be taken up in the order of their occurrence. Over half of them arranged topics with reference to another order. The preceding problem involved an arrangement of states as to their importance in a given industry. This "set" persisted. In another instance a pupil was asked to list the steps in a process. The work which had immediately preceded the question dealt with more general conceptions—namely, climate, location and soil condition. The answer given was in terms of the previous discussion.

Problems.—Some of the problems which grew out of these and similar experiences could profitably be made the subject of further study, and will, therefore, be stated in conclusion.

1. To what extent does "mental set" influence performance in standard subject tests?

2. Rapid mental adjustment is required by Binet Tests. If it is an intelligence factor which is affected by training as well as by age and development, how would general training in rapid mental adjustment affect Binet test scores?

3. To what extent must training of this ability be special training and to what extent does it transfer?

4. If drill devices are constructed from an analysis of functions into special habits, is it possible that an over training in special habits isolated for drill purposes contributes a mental "set" which must later be broken up?

5. What would lead pupils to put forth conscious effort in developing this ability?

6. To what extent can training involve a setting off of a group of reactions which is not peculiar to the special training itself, but which is a factor in other situations where that special habit must function?

PROBLEMS SUGGESTED FOR DISCUSSION IN SUCCEEDING ISSUES.

How Do Attitudes Affect School Progress?

Indications of the insufficiency of intelligence tests as prognostic devices. The possibility of improving attitudes and the relative constancy of the intelligence quotient. Consequent importance of the study of the effect of attitudes on school results.

What Can Be Learned From the Manifestations of Nervous Exhaustion in School Children?

Possible causes. What school experiences are apt to draw heavily on nerve force? How can the nervous equilibrium be re-established and maintained? Individual differences. Factors beyond school control.

What Are the Situations in Which Reading Functions?

The scope of training in the intermediate grades. Present tendencies as noted in new texts and studies. The next step.

What Is the Significance of the Varied Types of Imagery Used in Problem Solving?

Problem suggested by one test given to groups differing in age and grade. Possible effects on curriculum and learning of individuals.

Lincoln School of Teachers College. L. ZIRBES.

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES

Intelligence of Troops Infected With Hookworm vs. Those Not Infected. Garry C. Myers. Pedagogical Seminary, 1920, 27, 211-241. Those infected found to be of lower intelligence on the average than those not infected. The evidence does not prove that hookworm is the cause of inferior intelligence. Urges health education.

Intelligence at Senescence. M. F. Beeson. Journal Applied Psychology, 1920, December, 219-235. A study of the intelligence and education of inmates of public homes for the aged. An average I.Q. of 83 was found.

A Short Point Scale for Mental Measurement. Esther Reedy and James W. Bridges. Journal Applied Psychology, 1920, December, 258-262. Description of an abbreviated form of the Yerkes-Bridges Point Scale.

A Standardization and Weighting of Two Hundred Analogies. Rudolph Pintner and Samuel Renshaw. Journal Applied Psychology, 1920, December, 263-273.

Tentative Norms in the Rational Learning Test. Joseph Peterson. Journal Applied Psychology, 1920, December, 250-257.

The Accomplishment Quotient. Raymond Franzen. Teachers College Record, 1920, 21, 432-440. The theory and practice of basing school marks upon the quotient obtained by dividing the Educational Quotient by the Intelligence Quotient, thus displaying in terms of percentage the relation of actual progress to possible progress.

Table for Computing Mean Individual Scores in Educational Scales. Marvin J. Van Wagenen. Teachers College Record, 1920, 21, 441-451.

An Attempt to Measure the Comparative Importance of General Intelligence and Certain Character Traits in Contributing to Success in School. S. L. Pressey. Elementary School Journal, November, 1920, 220-229. Deals with health, "school attitude," preparation and ability as factors which condition success in school and success on a scale of intelligence.

Chicago Intelligence Test in Harrison Technical High School. O. Winter. School Review, December, 1920, 772-775. High School freshmen tested and grouped in teaching sections by the Chicago Group Intelligence Test. Moderate correlations shown between tests results and other measures of school ability like "marks."

Problem-Solving or Practice in Thinking (III). S. C. Parker. Elementary School Journal, November, 1920, 174-188. Third of a series of four articles

which consist of (1) account of the place of problem-solving in everyday life; (2) two sample lessons from the University of Chicago Elementary School; (3) "how great problem-solvers think."

Value of the Controlled Mental Summary As a Method of Studying. S. L. Pressey and L. W. Pressey. School and Society, November 27, 1920, 531-534.

Measuring the "Usefulness" of Tests in Solving School Problems. S. L. Pressey and L. W. Pressey. School and ociety, November 27, 1920, 531-534.

The Educability Level of Five-Year-Old Children. G. G. Ide. The Psychological Clinic, May, 1920, 13, p. 146-173. An intensive investigation of the mental abilities of two groups of kindergarten children with prognostic suggestions. Clinical histories of many cases are given.

Some Volititional Patterns Revealed by the Will-Profile. June E. Downey. Journal Experimental Psychology, 1920, 3, 281-302. A method of measuring certain temperamental rates with which measures of intelligence may be supplemented.

Clinical Method As a Method in Experimental Education. Frank N. Freeman. Journal Applied Psychology, 1920, December, 126-142. Some reflections upon the relationship between group and individual study as a method of educational investigation; an illustration of the individual study of a case of "word blindness' and a statement of some of the advantages of such a method.

"The Attempt to Teach." Gladys E. Poole. Psychological Clinic, May, 1920, p. 173-190. Diagnosis by means of mental and educational tests and the method of clinical teaching of five cases of backwardness in reading.

Esthetics. Rudolph Pintner. Psychological Bulletin, 1920, 17, 331-335. A summary of articles appearing during the last year. Includes articles on aesthetic appreciation of children, measures of musical talent, etc.

The Psychology of the Thrill. Irving R. Kaiser. Pedagogical Seminary, 1920, 27, 243-281. A theory that emotions, instincts, interests, will, etc., are but means whereby thrills are sought and experienced; and which, in being experienced, afford us temporary relief from the feelings of stress which are brought about by the repression which civilization has placed upon our more primitive desires and actions.

Opportunities for College Graduates in Psychological Examining in Social Service Work and Education. Mildred Francis Baxter. Journal Applied Psychology, 1920, December, 207-219. A survey of positions now open to women trained in psychology, with an estimate of future possibilities.

Racial Differences in Mental Fatigue. Thos. M. Garth. Journal Applied Psychology, 1920, December, 235-245. Indians are less susceptible to mental fatigue as tested than whites, who are less susceptible than negroes.

Work, Fatigue and Inhibition. F. C. Dockeray. Psychological Bulletin, 1920, 17, 322-330. Summaries of articles on the psychology and physiology of work, etc., which have appeared during the past year.

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION*

I. Books and Monographs in Applied and General Psychology.

The Psychology and Education of Subnormal Children: A Book for Teachers.—With the rapid extension of special classes for the education of subnormal children has developed an increasing demand for teachers who are specially trained both in the psychology and teaching of such children. A recent volume by Leta S. Hollingworth is written primarily for such teachers.[1] It gives an exceptionally good account of the "educational psychology" of subnormal children with the essentials of the physiological, medical and legal aspects of the subject. The causes of mental deficiency, the nature of individual differences, the means of identification, classification, definitions, the proportion of defectives among delinquents, prostitutes, unemployed, school laggards, etc., the physical traits, the nature of mental growth and maturity, special abilities and other features of feeblemindedness are adequately discussed.

Two important chapters deal with the causes and prevention of mental deficiency; the effects of heredity, alcoholism, toxins, sensory defects, tonsils and adenoids, malnutrition, hookworm, malasia, injuries to nervous tissues, syphilis, glandular defects and the like. Speech defects, hysteria, dementia praecox and other nervous and mental disorders which may complicate mental deficiency are briefly considered.

The book emphasizes the educational aspects of subnormality. The defective is found to be more like the normal child in physical size and strength, in sensory acuity and in motor control, hence the emphasis upon sensori-motor instruction. The defective is found to resemble the normal child in instincts and emotions. His impulses are strong, but the capacity to acquire control of them, weak. With the feebleminded, "moral training takes place only through the formation of specific habits." However, the author makes it clear that the learning of the feebleminded is not different in kind from the learning of the normal child of *the same mental age*. In general, the procedure is the same, the rate is the same,

*This department is conducted this month by A. I. Gates and H. O. Rugg.

[1]Leta S. Hollingworth. *The Psychology of Subnormal Children.* New York: Macmillan, 1920. Pp. XIX + 288.

the amount of transfer the same. It is pointed out, however, that differences in emotional reaction, in habits, kind of information found in two children, one feebleminded, one normal, both of mental age 8, are to be expected. Likewise their equal capacity to learn is only temporary since the growth, mentally, is different, hence the necessity for special classes or special schools, the history, organization and conduct of which are reviewed. The mental level ulutimately reached by the defective is subnormal, so different content to some extent, and different educational and vocational aims than those of the customary school are demanded. The book includes what is useful for the teacher of special classes or for the teacher of such teachers. Each chapter is followed by an excellent list of references. The material throughout is well organized and skillfully written.

Books Dealing With Mental Disorders and Mental Hygiene.—Books dealing with mental disorders, their diagnosis and treatment, especially from various branches of the Freudian School, have been flowing from the press within recent years. Already psycho-analysis is being urged as a means of solving educational problems by applications of the method to children. Jeliffe (The Technique of Psycho-analysis, 1918) states that "a new science and application of pedagogy are being reared upon the data obtained by psycho-analysis." The theories underlying most of their systems are in most respects not in harmony with the accepted principles of experimental and systematic psychology, but psychologists have been slow to print thoroughgoing evaluations of these doctrines. Dunlap[2] has given us a timely and critical evaluation of the outstanding theories and practices of the Freud-Jung-Adler schools. Psycho-analysis is the "most important mystical movement of the nineteenth century" and "the most significant feature of mysticism which comes strangely to the surface in the Freudian school is its antipathy to experimentation."

Chapter I presents a brief history and description of mysticism with which psycho-analysis is identified. Chapter II gives an excellent exposition of the important postulates of the several systems, with evidence of the continued appearance of logical fallacies, the disregard of accumulated scientific data, and an almost

---

[2]Knight Dunlap. *Mysticism, Freudianism and Scientific Psychology.* St. Louis: C. V. Mosby Co., 1920. Pp. 173.

complete disregard of the recognized methods of scientific procedure and proof. In the third chapter the psycho-analysists are given, it appears, some much needed first lessons in scientific methods followed by an exposition of some elementary facts concerning the nervous, muscular and glandular mechanisms, instincts, emotions, ideas, the laws of habit formation, etc. The book is apparently directed to the "general reader," but many pages will be specially useful to the student of education and psychology.

Destructive criticism is important, and just now much needed, in this field, but constructive contributions from the ranks of psychologists are even more to be desired. Hollingworth,[3] as a result of his experiences as director of the psychological service of the U. S. A., General Hospital No. 30, at Plattsburg, has given us the most complete study of certain types of the functional neurosis which has appeared. The work is important because the examinations in most cases were elaborate and because probably two-thirds of the group of the 1200 soldier patients presented persistent psychoneurotic symptoms which were not severe enough to lead them to seek treatment under conditions of ordinary civil life. The book discloses the methods and tools employed by the psychological service in diagnosis and remedial treatment, together with a summary of statistical findings. For the first time, measures of the general intelligence level of such cases are provided, the average mental age being 11.7 years, nearly two and a half years below that of the average soldier. These age distributions show a striking parallel with educational attainments; 26 per cent received no schooling, but 15 per cent reached the 8th grade and but 4.4 per cent graduated from High School. When the cases were grouped according to symptoms: Group I, including those whose symptoms were *physical*, i. e., fits, tremors, stuttering, seizures, paralyses, sensory disturbances, blindness, cardiac troubles, etc.; Group III, those whose symptoms were definitely *psychic* or *subjective*, such as fears, worries, anxieties, obsessions, psychasthenia, nightmares, emotional disturbances, aboulias, depressions, etc.; and Group II. those suffering from both types of symptoms, and doubtful cases with such symptoms as headache, insomnia, chills, fatigue and fainting, a hierarchy of mental levels was found. For Group I,

---

[3]H. L. Hollingworth. *The Psychology of the Functional Neuroses.* New York: Appleton, 1920. Pp. XIII + 259.

the average mental age was 10.9 years; Group II, 12.0, and Group III, 14.5. Differences in specific symptoms are directly related to differences in mental level; the dullest displaying somatic disorders chiefly, those of slightly higher level displaying more strictly neurasthenic traits, while in the upper range are found the more exclusively mental and psychasthenic symptoms.

The book includes a wealth of statistical information from the use of the Woodworth Psychoneurotic Inventory, the change of symptoms following the armistice, the "scattering" of mental abilities, army rank distributions, occupational and geographical distributions, and evaluations of tests and methods. The data are analyzed and out of them has developed a theory of the mechanisms involved in psychoneurotic conditions. It is the simple and well known mechanism of associative shifting or the "redintegrative mechanism" which has long been fundamental in our explanations of perception and other integrated reactions. ABCD, etc., elements of a complex situation, some or all of them, produce a complex response XYZ. Later the common element C, met among other stimuli, F.G.H., "redintegrates" XYZ and later H, associated with C in the second case, appearing alone with J.K.L. redintegrates the response again. The theory is discussed in detail and that it may be readily substituted for such mystical or pictorial concepts as "sphoning of the affect," "repression into the unconscious," "transfer of libido," etc., is suggested. The explanation here offered is simple enough to satisfy the rigid requirements of the law of parsimony; it merits careful consideration and test in the hands of professional workers. It promises to be the most important theoretical contribution in this field within recent years.

Significant details of two years' experience as an independent clinical psychologist are given in a brief account by Martin,[4] formerly Professor of Psychology at Stanford University. An outline of methods of diagnosis and treatment, largely educational, with representative case studies is given. The book contains little new theory, practice being based on the principles of the Freudian School and indirectly on the methods of experimental psychology, with a rather unusual reliance upon transfer effects. The psychologist may doubt the value of many of the remedial exercises that are suggested despite the appearance of cures. As Janet reminds

---

[4]Lillian J. Martin. *Mental Hygiene.* Baltimore: Warwick and York, 1920. Pp. VI + 89.

us, "The temple of Æsculapius has cured thousands of patients; Lourdes has cured thousands of patients; animal magnetism has cured thousands of patients; Christian Science has cured thousands of patients—and psycho-analysis has cured thousands of patients.—It is extremely difficult to eliminate other influences which may have modified the disease." Suggestion, stimulation, encouragement, sympathy, ridicule or other incidents within or outside the control of the practitioner, rather than the pictures, the gymnastics, the music or practice on control of imagery may have been the cause.

A conventional classification and statement of principles of causation and treatment of mental disorders will be found in a recent syllabus prepared by Charles B. Thompson.[5]

A New Manual of Mental Tests.—The scientific movement in education would be much enhanced if "directors of research," and designers of intelligence and educational tests would apply the type of careful procedure and cautious interpretation which the authors of the new Dewey-Child-Ruml test manual have employed. There is a good deal of reason to believe that we "scientific educationalists" are developing a rather crude and uncritical procedure—especially so, in our hasty publication of tests and test results; in our slip-shod making and testing of hypotheses (many of us never bother to make them in our investigations and studies) and in our naive interpretation of "correlations," "probable errors," forms of distribution and the like. This new book[6] will serve as a fine stimulus to more careful work in the testing field. These workers, leaders in a wider co-operative group, have tried out, assembled into a working unit, and have standardized in a preliminary way a number of new mental tests, together with groups of tests already in current use. Their purpose was to discover if an individual analysis (of normal pupils) can be made *by* tests, not *from* them: an analysis which is definitely the outcome of purely objective test results. In addition to using such available tests as the Binet, the Yerkes-Bridges Point Scale, the Woolley Identification of Forms, the Healey Picture Completion test and the like, the group (especially Miss Dewey and Professor Hayes) have devised a number of new tests. These include such examples as: cart construction, narrative pictures, needle threading, nail driving, etc.

---

[5]Charles B. Thompson. *Mental Disorders.* Baltimore: Warwick and York, 1920. Pp. 48.
[6]Dewey, E., Child, E., and Ruml, B. *Methods and Results of Testing School Children.* New York: E. P. Dutton & Company, 1920. Pp. XII + 176.

The testing was done with individual children, the time of testing varying from five to six periods, from twenty-five to forty-five minutes each. From these tests an abbreviated maturity scale was prepared, one for boys, another for girls. The one for boys consisted of seven tests; that for girls of six tests. In making this scale the writers avoided the assumptions which underly the different age-grade scales in current use—namely, "that a test situation gets at the essential factors in general intelligence and that tests highly correlated with their own total were necessary functions in identifying intelligence." They have so selected their tests that they do not overlap, and with reference to a measurable criterion *outside the test series.* They have stated the scores in terms of relation of each child's deviation from the average to the standard deviation of his age group.

Test scores and norms are reported in detail, together with diagrams showing the complete distribution of scores made on each test. The statistical methods employed are unique for such publications. For the first time definite use has been made of the regression equation in stating scores and norms. They report the regression of "scores on age" instead of using the commoner method of giving the arithmetic mean (or median). They did this so that erratic effects obtained by sampling would be neutralized.

We have held up this manual as a useful model of scientific pre cision in formulating and standardizing tests. (Its cautious interpretations also set an example in the treatment of test material.) We would emphasize the fact that this is a useful example of scientific technique and *nothing more.* We would not have it thought that we believe that the procedure and results can be employed in public schools today. They cannot. They are administratively impossible. At the present time probably not twenty-five educationalists in America can interpret and use its procedure and results correctly without further careful statistical training. It is possible, however, to obtain from this little book (impracticable as it may be today) an object lesson for the improvement of our investigational technique and of our own mental attitudes toward rigorous scientific work in education.

The Eyesight of School Children. A treatment of the problem of vision in the schools of an authoritative character and more complete than the conventional chapter in books on hygiene, is needed.

Berkowitz[7] has fulfilled this need by presenting a summary of facts concerning prevalent eye defects with a series of recommendations concerning diagnosis and treatment which have been approved by a committee of leading oculists and ophthalmologists. The relation of day and artificial lighting systems, the arrangement of desks, interior decorations, blackboards, etc., to the hygiene of vision are samples of problems which are treated.

II. BOOKS AND REPORTS IN EDUCATION.

Two New Books on the Elementary School Curriculum.—Two books have appeared recently which deal with the reorganization of the elementary school curriculum.[8, 9] The writer of each one has spent some years in the reorganization of school curricula. Each writer is applying the best there is in the James-Dewey-Thorndike pragmatic and social philosophy and psychology. While both books plead for "reorganization," yet the two are nearly as unlike as it is possible for them to be.

Professor Bonser agrees with Professor Meriam that it is "ideally desirable" *to organize a curriculum about activities* rather than subjects; that people's activities shall be so utilized that "the curriculum shall directly emerge *from* them and hence, fit *for* them." The curriculum emerges as "a series of purposeful activities," and is the "means of providing them." Each of these two books aims to do what our pragmatic and social philosophy has been preaching for a generation: that is, construct a curriculum completely on the criterion of social worth. "Only as the knowledge, habits and attitudes, and appreciations developed in the school are operative in meeting the problems of life are they of any worth." To this statement they will both subscribe.

Striking differences appear in the applications which the two men make of their theories. While it is "ideally desirable," yet Professor Bonser thinks it is "practically" impossible to completely organize a curriculum about "activities." His method is to lead school men gradually to change their procedure. So we find, in his book, a discussion of all the established school subjects. His device

---

[7]J. H. Berkowitz. *The Eyesight of School Children.* Washington; Govt. Printing Office, Bureau of Education, Bulletin No. 65, 1920. Pp. V + 128.

[8]Bonser, F. G. *The Elementary School Curriculum.* New York: Macmillan, 1920. Pp. XVI + 446.

[9]Meriam, J. L., *Child Life and the Curriculum.* Yonkers: World Book Co., 1920. Pp. XII + 538.

for effecting change gradually is the "project." "Changes must be gradual, not abrupt for them to be adopted by teachers," untrained and yet habituated as they are to a "subject" basis of organization. He recommends a "natural transition," however, through the pages of his book. "Beginnings may be made with single units," this means blocks of material organized as problem-projects.

Professor Meriam's book is, first, a more vigorous denunciation of the existing curriculum. Several chapters are devoted to this sort of argument and to developing principles of curriculum-making.

When we come to his "contents of a curriculum" we find that Professor Meriam has indeed had the courage of his convictions and has completely broken away from the established school subjects.

He sets up four school studies in his school (The University Elementary School, Columbia, Mo.): 1—Observation; 2—Play; 3—Stories; 4—Handwork. Here the program is *not* organized as reading, writing and arithmetic; history, geography, science, civics are all excluded from his curriculum scheme. To be sure the material of these "studies" is all classifiable under one or more of the established subjects. Quite clearly the writer confuses *general* activities (like play) with devices for stimulating such (like stories), and in turn with mental functions (like observation) by discussing these as though they were coordinate. Observation is to form the core of the work in social and industrial activities of the intermediate grades, animal life, earth and sky and people in the primary grades; local industries in grades III and IV. Stories including music, poetry and pictures is the one phase of the four-fold scheme that is to train primarily for the *leisure* life, according to the writer. The burden of his argument is, first, that by organizing "reading" as an incidental school activity reading becomes enjoyable and that the amount read is large. No definite measurements are reported to show to what extent the children in the school actually *can* read, however. This is *true* throughout the book.

The book is full of inconsistencies and contradictions; for example, its frequent questioning of the worth of quantitative methods, while subsequently calling to the author's support just such methods. It is unorderly, very repetitious, and it appears to the reviewer, needlessly argumentative and exhortatory. It is clear that it could not be used as a text for *prospective* teachers—it would confuse rather than help. For the experienced and capable teacher,

for trained faculties of normal schools and city school systems and for graduate students in education, we believe Professor Meriam's book will serve as a "stimulating irritant" to the discussion of the mooted questions of curriculum-making.

Professor Bonser's book on the other hand will serve as an admirable text book for prospective teachers or for reading-circle work. It is orderly, systematic, carefully written; will introduce students and teachers to modern problems and methods of organizing a curriculum and will show how the traditionalism of current practice may gradually be broken through. It will train teachers well for present jobs. The present reviewer doubts if it will serve to stimulate as much thoroughgoing discussion of the *reconstruction* of the curriculum, however, as will "The Child and the Curriculum," even though the latter may seem to some of us a definite example of what *not* to do.

A "Text"-Book on the Supervision of Instruction.—Books on the supervision of teaching have rarely contributed to improve the actual class-room work of teachers. The chief reason for this is that they have been compilations of platitudes about the purposes of teaching, sets of principles of "methods," general directions, etc. While Professor Nutt's new book [10] is an improvement on the mass of such books it cannot be said to set a really new mark in this field. It continues the precedent set by the earlier books, of giving many chapters of "Principles Underlying the Supervision of Instruction" (Chs. III, IV, V and VI), and scores of pages of "devices of supervision," which are largely *comment about devices* and not the devices themselves. Occasionally one meets concrete *examples* of supervision (e. g. ps. 158-59); now and then a "device" is presented in sufficient detail to inspire a supervisor to go and do likewise. One reads the pages of this book, however, with dismay at the accumulation of principles and methods and with an unsatisfied sense of not having seen and learned how to use the devices and methods which good supervision must be employing. Its chief function will be to serve as a guide for college class discussions in which the instructor supplies the rich background that is needed for understanding the "texts" of the book.

Its principles of method and its survey of devices are in sound accord with our current psychology. The suggestions for the actual

[10] Nutt, H. W. *The Supervision of Instruction.* Boston: Houghton Mifflin Company, 1920. Pp. XI + 277. $1.80.

supervision of class work are sensible and helpful. Its great deficiency is in omitting a wealth of pertinent *examples* for each phase of its discussion.

The Need of Federal Assistance for the Nation's Schools.—Since "the entire schooling of the average native-born citizen has cost the public less than one hundred and fifty dollars—an amount of comparable perhaps with that which the village grocer spends on his daughter's piano lessons," and since it is found that in five prosperous Middle-Western States the machinist, plumber, bricklayer and blacksmith are paid more than twice the average salary of elementary teachers, Keith and Bagley,[11] for these and other reasons, have advocated the creation of a Department of Education provided with means by the Federal Government for remedying the situation. Aside from a comprehensive statistical study of the existing conditions among schools with an effort to trace them to their causes, the book outlines the historical development of the policy of Federal aid, eventuating in a critical consideration of the Smith-Towner Bill, as representing the most satisfactory current proposal. Although the book is written in form suitable for the general reader, it cannot be too strongly emphasized that the professional man in education and psychology should seek intimate acquaintance with material such as is here assembled.

A Committee Report on the Reorganization of Science in Secondary Schools.—After more than seven years of discussion and experimentation, the Committee on Reorganization of Science in the Secondary Schools, under the leadership of Professor Caldwell,[12] has given a report which is a marked improvement upon the conventional "writ of opinion" variety. Although no experimental data are given, the suggestions with regard to materials and methods are offered after having been successfully tried in many schools. Projects, of a sane and useful sort, carried out under careful guidance by the teacher is the foundation of the method. The relation of projects to topics, "the natural way of working," choice of topics and materials, laboratory and classroom procedure, demonstrations, excursions and clubs are features which are considered in detail. While this sort of work may be in no sense considered the equivalent

---

[11]John A. H. Keith and William C. Bagley. *The Nation and the Schools.* New York: Macmillan, 1920. Pp. XVII — 364.

[12]Otis W. Caldwell and others. *Reorganization of Science in Secondary Schools.* Washington: Govt. Printing Office, Bureau of Education, Bulletin No. 26, 1920. Pp. 62.

of experimental research and experimental teaching, it marks a distinct improvement in committee procedure.

III. Additional Publications Received.

A. Publications in general and applied psychology.

Edman, I. *Human Traits and Their Social Significance.* Boston: Houghton Mifflin Company, 1920. Pp. XI + 467.

Godin, P. *Growth During School Age.* Boston: Badger, 1920. Pp. 268. Translated by S. L. Eby.

Lay, W. *Man's Unconscious Passion.* New York: Dodd, Mead & Co., 1920. Pp. V + 245.

Muscio, B. *Lectures on Industrial Psychology.* London and New York: George Routledge & Sons, Ltd., and E. P. Dutton & Co., 1920. Pp. 1 + 300.

Schofield, A. T. *The Mind of a Woman.* New York: E. P. Dutton & Co., 1919. Pp. VII + 120.

Seashore, C. E. *A Survey of Musical Talent in the Public Schools.* Iowa City: The University, Iowa City, Ia., 1920. Pp. 7 + 36.

Terman, L. M. *Condensed Guide for the Stanford Revision of the Binet-Simon Intelligence Tests.* Boston: Houghton Mifflin Company, 1920. Pp. 7 + 32.

Williams, J. H. *A Survey of Pupils in the Schools of Bakersfield, California.* Whittier, Cal.: Whittier State School, Department of Printing Instruction, 1920. Pp. 1 + 43.

B. Publications in the general educational field.

Briggs, T. H. *The Junior High School.* Boston: Houghton Mifflin Company, 1920. Pp. IX + 350.

Cubberley, E. P. *The History of Education.* Boston: Houghton Mifflin Company, 1920. Pp. XI + 849.

Cubberley, E. P. *Readings in the History of Education.* Boston: Houghton Mifflin Company, 1920. Pp. IX + 684.

Hanus, P. H. *School Administration and School Reports.* Boston: Houghton Mifflin Company, 1920. Pp. III + 200.

Showalter, N. D. *A Handbook for Rural School Officers.* Boston: Houghton Mifflin Company, 1920. Pp. IX + 213.

Stockton, J. L. *Project Work in Education.* Boston: Houghton Mifflin Company, 1920. Pp. V + 166.

C. Miscellaneous publications.

Findlay, J. J. *An Introduction to Sociology.* London: Longmans, Green & Co., 1920. Pp. IX + 304.

Lambertson, F. W. *The Rules of the Game.* New York: The Abingdon Press, 1920. Pp. 5 + 208.

Lambertson, F. W. *The Rules of the Game. Teachers' Manual.* New York: The Abingdon Press, 1920. Pp. 5 + 77.

Parker, D. H. *The Principles of Aesthetics.* Boston: Silver, Burdett & Co., 1920. Pp. V + 374.